

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky  
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA ELEKTROTECHNIKY  
A KOMUNIKAČNÍCH TECHNOLOGIÍ**

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

**ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ**

DEPARTMENT OF BIOMEDICAL ENGINEERING

**EPIDEMIOLOGICKÉ MODELOVÁNÍ ŠÍŘENÍ NEMOCI  
COVID-19**

EPIDEMIOLOGICAL MODELING OF COVID-19

**BAKALÁŘSKÁ PRÁCE**

BACHELOR'S THESIS

**AUTOR PRÁCE**

AUTHOR

Richard Schubert

**VEDOUCÍ PRÁCE**

SUPERVISOR

Ing. Martin Mézl, Ph.D.

**BRNO 2021**

# Bakalářská práce

bakalářský studijní program **Biomedicínská technika a bioinformatika**

Ústav biomedicínského inženýrství

**Student:** Richard Schubert

**ID:** 195717

**Ročník:** 3

**Akademický rok:** 2020/21

**NÁZEV TÉMATU:**

## Epidemiologické modelování šíření nemoci Covid-19

### POKYNY PRO VYPRACOVÁNÍ:

1) Prostudujte problematiku epidemiologických modelů a proveďte rešerši základních i pokročilých modelů. 2) Proveďte rešerši využití epidemiologických modelů při retrospektivní analýze epidemických dat pro onemocnění Covid-19. 3) Proveďte simulaci základních epidemiologických modelů a jejich srovnání. 4) Na datech z veřejných databází proveďte odhad parametrů zvoleného epidemiologického modelu. 5) Modelujte základní vliv protiepidemických opatření a snažte se jej interpretovat v dostupných datech. 6) Proveďte diskuzi získaných výsledků. Zaměřte se na věrohodnost modelů pro různé průběhy epidemií a jejich vzájemné srovnání.

### DOPORUČENÁ LITERATURA:

[1] HE, Shaobo, Yuexi PENG a Kehui SUN. SEIR modeling of the COVID-19 and its dynamics. Nonlinear Dynamics. DOI: 10.1007/s11071-020-05743-y. ISSN 0924-090X.

[2] ALAMO, Teodoro, Daniel REINA, Martina MAMMARELLA a Alberto ABELLA. Covid-19: Open-Data Resources for Monitoring, Modeling, and Forecasting the Epidemic. Electronics. 2020, 9(5). DOI: 10.3390/electronics9050827. ISSN 2079-9292.

**Termín zadání:** 8.2.2021

**Termín odevzdání:** 28.5.2021

**Vedoucí práce:** Ing. Martin Mézl, Ph.D.

**doc. Ing. Jana Kolářová, Ph.D.**  
předseda rady studijního programu

### UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

## ABSTRAKT

Tato práce se zabývá spojitými epidemiologickými deterministickými kompartmentovými modely a specifickými rysy modelování pandemie COVID-19. Numericky je studována rozdílnost různých pravděpodobnostních rozdělení setrvání jedinců v kompartmentech na základní reprodukční číslo, respektive konečnou velikost epidemie. Je navržen nový model pro retrospektivní analýzu epidemiologických dat ze severní Itálie první poloviny roku 2020. Odhad parametrů modelu je proveden minimalizací sumy váhovaných čtvercových reziduí a prohledávání prostoru parametrů implementací algoritmu BFGS.

## KLÍČOVÁ SLOVA

SARS-CoV-2 , COVID-19, Nelineární dynamika, Matematická epidemiologie, SEIR model, gama rozdělení, BFGS algoritmus

## ABSTRACT

This thesis deals with the continuous epidemiological deterministic compartmental models and the COVID-19 pandemic modeling distinctive features. The effect of different probability distributions of individuals stay in compartments is studied numerically in relation to basic reproductive number and the final size of the epidemic, respectively. New model for a retrospective analysis of the first half of 2020 northern Italy epidemiological data is proposed. The model parameters estimation is performed using minimisation of weighted sum of squared residuals and the search through parameter space with BFGS algorithm implementation.

## KEYWORDS

SARS-CoV-2 , COVID-19, nonlinear dynamics, mathematical epidemiology, SEIR model, gamma distribution, BFGS algorithm

## Prohlášení autora o původnosti díla

**Jméno a příjmení autora:** Richard Schubert  
**VUT ID autora:** 195717  
**Typ práce:** Bakalářská práce  
**Akademický rok:** 2020/21  
**Téma závěrečné práce:** Epidemiologické modelování šíření nemoci Covid-19

Prohlašuji, že svou závěrečnou práci jsem vypracoval samostatně pod vedením vedoucí/ho závěrečné práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno .....

.....

podpis autora\*

---

\*Autor podepisuje pouze v tištěné verzi.

## PODĚKOVÁNÍ

Rád bych poděkoval panu Ing. Martinu Mězlovi, Ph.D., který mě po celou dobu, kdy tento text vznikl, odborně vedl.

# Obsah

Úvod	8
<b>1 Matematické modely v epidemiologii</b>	<b>11</b>
1.1 SIS model . . . . .	11
1.2 SIR model . . . . .	12
1.3 SEIR model . . . . .	16
1.4 Modelování karantény a izolace . . . . .	18
1.5 Gama distribuce infekčnosti . . . . .	20
1.6 Počátek epidemie a <i>network based</i> modely . . . . .	22
1.7 Komparace vybraných modelů . . . . .	27
<b>2 Rešerše modelování COVID-19</b>	<b>32</b>
<b>3 Design modelu a simulace</b>	<b>38</b>
3.1 Model . . . . .	38
3.2 Data . . . . .	42
3.3 Odhad parametrů . . . . .	45
3.4 Výsledky a diskuze . . . . .	47
<b>Závěr</b>	<b>53</b>
<b>Literatura</b>	<b>54</b>

# Seznam obrázků

1.1	Některé průběhy rov. (1.4)	13
1.2	Diagram modelu SEQIHR	19
1.3	Porovnání gama a exponenciální distribuce	21
1.4	Grafická interpretace existence kořenů $z = G_1(z)$ .	24
1.5	Rozdíl v $\mathcal{R}_c$ pro gama a exponenciální distribuce	28
1.6	Efekt načasování kontrolních opatření	29
1.7	Ciudad de México 2009 - H1N1	31
2.1	Model podle [18] a simulace	33
2.2	Kolektivní imunita a průběh epidemie	36
3.1	Diagram modelu (3.1)	39
3.2	Analýza dynamiky (3.1) a senzitivity	41
3.3	Pandemie COVID-19 v severní Itálii 2020	43
3.4	Ukázka konvergence MAE pro algoritmus BFGS	46
3.5	Model vs data regiony severní Itálie	49
3.6	Model vs data Itálie	50
3.7	Hypotetický vliv absence kontrolních opatření	52



# Úvod

Pandemie infekční nemoci COVID-19 je dosud stále probíhající nákazou, které zasáhla většinu zemí světa bez ohledu na úroveň jejich zdravotnických systémů a ochrany zdraví. Bezprecedentnost nastalé situace v globalizovaném světě vedla k podcenění infekčnosti nemoci a příliš pomalé reakci exekutivních orgánů jednotlivých států. V mnohých lokalitách došlo ke zhroucení zdravotní péče a zdravotnická zařízení musely přejít do nouzového režimu, který zahrnoval odložení neakutních zákroků a přerozdělení personálu z jednotlivých specializací na jednotky intenzivní péče. Již z počátku pandemie bylo zřejmé, že nutné restriktce budou mít závažné ekonomické důsledky a světové hospodářství upadne do recese či ekonomické deprese.

Leitmotivem této práce je analyzovat modely, jenž byly použity, respektive mohly potenciálně být při odhadu chování této epidemie v blízké budoucnosti popřípadě retrospektivně. Pivotem tohoto textu je srovnání těchto modelů. Tedy účelem nebude provést predikce ani návrhy strategií pro probíhající pandemii. V první kapitole budou představeny spojitě deterministické kompartmentové modely a minoritně modely se základem v teorii grafů, aby byla zachycena limitace deterministických modelů a vyjádřena podmínka ekvivalence s grafovými modely. Budou například zkoumány podmínky asymptoticky stabilních řešení apod. Druhá kapitola bude obsahovat shrnutí některých poznatků o modelování probíhající pandemie v literatuře. Na tomto základě bude ve třetí kapitole sestaven deterministický model a proveden numerický odhad parametrů tohoto modelu na COVID-19 územní kazuistice ze severní Itálie, tedy na základě dostupných dat, a simulaci protiepidemických opatření, jenž ovlivňují parametry modelu, a kvantitativní zhodnocení získaných výsledků. Všechny numerické simulace kompartmentových modelů budou provedeny metodou Runge–Kutta čtvrtého řádu.

Od neolitické revoluce a postupnému nárůstu velikosti a koncentrace populace bylo lidské společenství náchylné k často velmi decimujícím epidemiím komunitně přenosných nemocí, které i za velmi krátký čas dokázaly zredukovat stávající populaci. To mělo často za následek nejen ekonomické, ale i geopolitické změny. Například při rozšíření bubonického moru z dnešních italských přístavů ve 14. století na celý kontinent podlehl nemoci téměř jedna třetina populace celé tehdejší Evropy. Mnohé infekční choroby si zachovaly endemičnost či oscilační charakter, který mohl mít periodu i několik let. Jako příklad uveďme spalničky, které se vracely v západní Evropě a Nové Anglii v intervalu přibližně tří let od 19. století a jejichž šíření bylo zastaveno až vakcinací v šedesátých letech, nebo tuberkulózu, jejíž přítomnost je doložitelná endemicky v populaci od antiky bez jakékoliv prodlevy až do dnešních dnů [1]. (V Evropě došlo ke snížení incidence na několik stovek ročně díky vaki-

naci. Nové případy jsou často spojeny s migrací z populací, kde je incidence stále vysoká). Pokroky v mikrobiologii v 19. století umožnily popsání mnohých smrtících agens, ale také zavést systematickou antiseptickou ochranu pro pacienty a zdravotnický personál. Posledními pandemiemi, které měly podobnou škálu rozsahu jako COVID-19, můžeme srovnávat s propuknutím epidemie španělské chřipky po první světové válce (avšak ne co do letality) či dnes ještě stále enigmatickou chorobu *encephalitis lethargica* probíhající celosvětově v letech 1915-1926, postihující oběti velmi těžkým parkinsonismem, nebo Hongkongskou chřipkou také s globální incidencí v letech 1968–1970 [1].

Lékařský výzkum hrál a hraje nezastupitelnou roli v potlačování a profylaxi výše zmíněných chorob. Na rozdíl od geneticky podmíněných či různých civilizačních chorob nemohou být komunitně přenášené nemoci pro své bezprostřední ohrožení společnosti přenechány k eventuální budoucí diagnostice. I když dojde k identifikaci patogenu, je vždy nutné vypracovat model, který bude odrážet chování jedinců a jejich kontakt, který je nezbytný k šíření nákazy. Aby bylo možné k tomuto cíli dospět, musí být problém přenesen do abstraktního prostoru, kde dojde k redukci modelu šíření, jinými slovy opomenutí proměnných, které alespoň zdánlivě jsou nepodstatné pro predikci, neboť při zmenšující se abstrakci modelu dochází k nárůstu komplexnosti, která zahrnuje obtížnost nalezení řešení pro daný model, respektive interpretaci výsledků z modelu vyplývajících. Jinými slovy pro modelování infekčních nemocí je vždy nutné rozumět základní biologii choroby, poté tento problém vyjádřit v matematickém rozhraní, kde uvažuje pouze důležité rysy přenosnosti. Po vytvoření určitého předpokladu z výsledku modelu, je poté možné tento předpoklad přeložit zpět do biologie nemoci [2].

Základním cílem modelování při šíření nákazy je odhad úspěšnosti strategie při zvládnání epidemie, srovnání s jinými postupy a poskytnutí dat zdravotnickým a kontrolním orgánům pro vytváření ochranných opatření. Průkopnickou práci při aplikaci tohoto přístupu provedl *Ronald Ross* (druhá Nobelova cena za fyziologii a medicínu), který demonstroval mechanismus přenosu malárie mezi vektorem, komárem rodu *Anopheles*, a člověkem. Byl si vědom, že není možné dosáhnout absolutní extinkce komára v zasažených lokalitách, avšak matematickým modelem dovodil, že je pouze nutné udržet populaci komárů pod kritickou úrovní, aby došlo k zastavení epidemie. Tento závěr byl později potvrzen z terénních dat [2]. Myšlenka existence určité kritické hodnoty infekčnosti (přenosnosti), kdy epidemie zanikne či bude stále v určité míře pokračovat, je fundamentální rys pro analýzu epidemiologických modelů. Některé predikce modelů mohou být kontraintuitivní. Například je možné brát za zřejmé, že léčba infekční choroby by měla snížit její prevalenci, tj. poměr počtu infikovaných vůči všem jedincům v populaci v daném čase. Avšak existují situace, kdy dostupnost léčby může naopak podnítit vznik rezistentního kmene patogenu, tedy

vyšší frekvence jedinců kteří podstoupí léčbu ve skutečnosti zvýší incidenci choroby v populaci. Tato skutečnost je známým problémem z historie léčby tuberkulózy či AIDS. Modelování je kruciólní prostředek pro identifikaci těchto možných scénářů.

Použití matematických modelů v epidemiologii podnítily nejdříve lékaři v 19. století. Nejznámějším příkladem je práce *Johna Snowa* při identifikaci zdroje nákazy cholery v Londýně roku 1854 jako studny na Broad Street a to na základě počtu zemřelých na jednotlivé domy v nejvíce zasažených lokalitách. Již sofistikované modely byly základem při vytváření protiepidemických opatření ve 21. století zejména u epidemií slintavky a kulhavky ve Velké Británii roku 2001 (žádný přenos na člověka, hlavně ekonomické důsledky) a SARS epidemii 2002-2003 v Číně, Hongkongu a na Tchaj-wanu. Příznačné pro ověřování efektivnosti epidemiologických modelů je nemožnost testování modelů, neboť získávání dat při kontrolovaném experimentu by bylo neetické, tedy je nutné se spoléhat na data (neúplná) z již proběhlých epidemií. (Vděčným příkladem se stalo modelování šíření spalniček, neboť jde o onemocnění téměř jistě příznakové s typicky terčíkovitými vyrážkami a tedy je velice pravděpodobné zachycení incidence obvodním lékařem, tedy víme, že data odpovídají realitě.) Rigorózní přístup k modelování v epidemiologii (společně s modely v populační biologii) byl postupně budován od 30. let dvacátého století, avšak stále se jedná o okrajovou část aplikované matematiky, neboť všechny teoretické vlastnosti jsou odvoditelné z teorie dynamických systémů. Výsledkem výzkumu byly především strategické modely, které se spíše zaměřovaly na teoretické porozumění a generalizaci. V posledních desetiletích se od tohoto směru se z popudu lékařských profesionálů oddělil trend směřující k taktickému modelování, které se snaží osadit model, co nejvíce informacemi relevantních pro šíření choroby. Tedy dochází ke snížení abstrakce modelu. Tento přístup je možný díky pokrokům ve výpočetní technice, poskytující numerické řešení. Není zde důležitá stabilita ani existence analytického řešení, avšak pouze prediktivní schopnost [2].

Je nutné říci, že v modelování nebudeme pro zjednodušení brát na zřetel dlouhodobé demografické proměnné jako narození a smrt, případně budeme předpokládat, že jsou si přibližně rovny, či pokud nebude řečeno jinak. Jinými slovy fluktuace ve změně velikosti populace jsou škálově zanedbatelné vůči její celkové velikosti. Tyto proměnné vstupují do výpočtů v případě studia endemických onemocnění či predikcí na mnohem větší časové škále. Jak bylo řečeno, budou představeny hlavně spojité modely, pro něž však existují diskrétní analogie, jejichž predikce mohou být více úspěšné neboť data o incidencích choroby vždy přichází po diskrétních časových krocích. Avšak obecně jsou podmínky stability spojitých a diskrétních systémů odlišné, tedy předpovědi se mohou signifikantně lišit pro stejné počáteční podmínky.

# 1 Matematické modely v epidemiologii

Základem pro použití matematických modelů u infekčních nemocí se stala práce *W.O. Kermacka* a *A.G. McKendricka* vydaných ve třech článcích z let 1927-1933. Navrhli rozdělit populaci na kompartmenty, které jsou identifikovány podle stavu jedinců (infikovaní apod.) [3]. Model je zde vyjádřen jako systém obyčejných nelineárních diferenciálních rovnic. To implikuje plně deterministický charakter těchto modelů. Nejjednodušší případ, který Kermack a McKendrick uvažovali, bylo rozdělení populace na náchylné k infekci  $S(t)$  (susceptibles), respektive infikované  $I(t)$  v čase  $t$ . Avšak infikovaní ať představují pouze období choroby, kdy je jedinec infekční. Předpokládali jednoduchou závislost velikosti populace  $N$  a počtu kontaktů, které průměrný jedinec učiní podle zákona hromadné akce (*mass action incidence*). Kontakt je vždy efektivní (předpokládejme pro tento případ), tj. vždy dojde k přenosu infekce. Necht se tedy jedinec účastní v průměru  $\beta N$  sociálních styků za jednotku času v populaci o velikosti  $N$ , kde v tomto případě  $\beta$  považujeme za konstantu. Z tohoto počtu je průměrně  $I/N$  kontaktů s nemocnými. Nárůst infikovaných za jednotku času je poté možné vyjádřit jako  $\beta IS$ , neboť právě  $S$  jedinců činí kontakt, který způsobuje nárůst incidence. Můžeme tedy definovat jednoduchý dvoukompartimentový model [4].

## 1.1 SIS model

$$\begin{aligned} S' &= -\beta SI + \alpha I \\ I' &= \beta SI - \alpha I \end{aligned} \tag{1.1}$$

Tento model předpokládá extrémní zjednodušení. Jak je patrné, jedinci putují pouze mezi dvěma skupinami. Velikost populace je vyjádřena jako  $N = S + I$ . Sečtením obou rovnic obdržíme  $N' = 0$ , tedy neexistuje migrace do ani z populace, také infekce nemůže být smrtelná. Právý význam koeficientu  $\alpha$  vyplývá z druhé rovnice. Pokud uvažujeme pouze nemocné jedince v čase  $t = 0$ , potom počet jedinců, který zůstane v tomto kompartmentu přesně po dobu  $t_1$ , je dán  $-I'(t_1)$ . Vyjádření průměrné doby v kompartmentu je  $\int_0^\infty -tI'(t) dt$  a poté například přes *per partes* dostáváme hodnotu  $1/\alpha$ . Tedy jedinec je nemocný průměrně  $1/\alpha$  času.

Tento model je jako jeden z mála plně analyticky řešitelný. Dosazením  $S = N - I$  do druhé rovnice (1.1) dostáváme úpravu

$$I' = \beta I(N - I) - \alpha I \implies I' = rI \left(1 - \frac{I}{K}\right). \tag{1.2}$$

Pro zjednodušení jsme zavedli  $r = \beta N - \alpha$ ,  $K = N - \alpha/\beta$  (v populační biologii *carrying capacity*). Jedná se o logistickou rovnici. Tu je možné vyřešit přímo

separací proměnných nebo elegantněji zavedením substituce funkce  $I(t) = 1/z(t)$ , která rovnici (1.2) redukuje na obyčejnou lineární diferenciální rovnici prvního řádu. Například přes integrační prvek a následnou zpětnou substituci obdržíme

$$I(t) = \frac{KI_0}{I_0 + (K - I_0)e^{-rt}}, \quad (1.3)$$

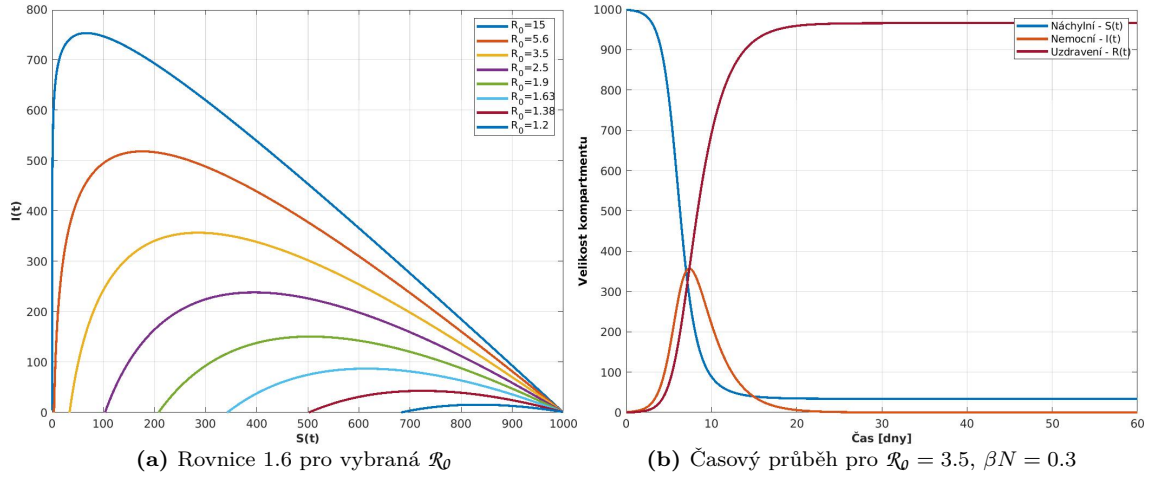
kde  $I_0$  značí počáteční počet infikovaných. V případě  $r > 0$  vidíme globální asymptotickou stabilitu pro všechny  $I_0 > 0$ , kdy  $I \rightarrow K$  pro  $t \rightarrow \infty$ . V opačném případě  $r < 0$  dochází pro  $K > I_0 > 0$  ke konvergenci  $I \rightarrow 0$ . Tento aspekt má velký epidemiologický význam, neboť jiným přepsáním  $r$  získáme  $\beta N/\alpha$ . Tento poměr nazveme *reprodukční číslo* se značením  $\mathcal{R}_0$ . Existuje omnipresence takového čísla v určité podobě ve všech kompartmentových modelech (podobně jako Rayleighovo číslo ve fluidní mechanice), avšak mnohdy s mnohem složitějším dovozením. Funguje jako práh, který určuje zda k propuknutí epidemie dojde či nikoli. Pro SIS je zřejmě  $\mathcal{R}_0 = \beta N/\alpha > 1$ , jenž způsobuje přetrvání epidemie (konvergence ke  $K$ ), naopak  $\mathcal{R}_0 < 1$  zaručuje vymýcení choroby.

$\mathcal{R}_0$  může být v tomto případě interpretováno jako množství sekundárních infekcí, které způsobí jeden průměrný jedinec během trvání celé epidemie. Podle tohoto modelu a dalších zde představených, pokud by na počátku epidemie platilo  $\mathcal{R}_0 < 1$ , k propuknutí šíření vůbec nedojde, jak je například zřejmé z druhé rovnice (1.1). Chování řešení je také určeno počáteční hodnotou  $I_0$ . Kdyby platilo  $I_0 > K$ , nemohlo by řešení nikdy poklesnout pod  $K$ . Neboť  $K$  tvoří kritický bod rovnice (1.3) (teoretické dovození například v [7]).

SIS model byl zmíněn kvůli dynamice, neboť již zde lze pozorovat dvě ekvilibria (pro  $\mathcal{R}_0 < 1$ ), i když není epidemiologicky realistický, neboť dochází k okamžitému návratu jedinců do třídy  $S$  po pominutí infekčnosti. Jedná se však o nejjednodušší model, kde existuje endemické ekvilibrium. Avšak lze se setkat s využitím (1.3) například jako logistické regrese pro určení kategoricky závislých proměnných či úpravy vztahů mezi více proměnnými modelu, či dále je přeformulovaný model (1.1) používán v takzvaném Lagrangianském rámci (*Lagrangian framework*), kdy byl například aplikován k vyhodnocení vlivu pohybu obyvatelstva při přenosu viru Eboly nebo viru Zika [6].

## 1.2 SIR model

První model, který lze považovat za epidemiologicky věrohodný, může být přidání třídy  $R$ , jenž představuje vyléčené jedince (*removed*), kteří jsou ovšem již perma-



**Obr. 1.1:** Některé průběhy rov. (1.4) pro  $N = 1000$  a  $I_0 = 1$ .

nentně imunizování [4]

$$\begin{aligned} S' &= -\beta SI \\ I' &= \beta SI - \alpha I \\ R' &= \alpha I \end{aligned} \quad (1.4)$$

Předpokládáme  $R_0 = 0$  a  $I_0 + S_0 = N$ , tedy můžeme třetí rovnici vynechat, neboť je plně určena zbylými dvěma. Řešení je možné nalézt následující úvahou. Sečtením prvních dvou rovnic (1.4) obdržíme

$$(S + I)' = -\alpha I \quad (1.5)$$

Z první rovnice (1.4) vidíme, že pro  $S_0 > 0$  je  $S(t)$  klesající funkce pro všechna  $t$ . Tvrdím, že musí platit  $\lim_{t \rightarrow \infty} -\alpha I(t) = 0$ . Kdyby nikoliv, tak předpokládejme  $\lim_{t \rightarrow \infty} I(t) = c > 0$ . Existuje okolí  $\varepsilon$  a  $t_0$ , kdy pro  $t > t_0$  platí  $|I(t) - c| < \varepsilon$ . Necht' je  $\varepsilon = c/2$ . Potom je ovšem (1.5) pro  $t > t_0$  monotónně klesající a tedy muselo by existovat  $t_1$  takové, že  $S(t_1) + I(t_1) = 0$ , což ovšem podle předpokladu není možné.

Integrací (1.5) dostáváme

$$\alpha \int_0^\infty I(t) dt = S_0 + I_0 - S_\infty = N - S_\infty. \quad (1.6)$$

Vydělením první rovnice (1.4)  $S$  a kombinací s rovnicí výše dává

$$\ln \left( \frac{S_0}{S_\infty} \right) = \beta \int_0^\infty I(t) dt = \frac{\beta}{\alpha} [N - S_\infty] = \mathcal{R}_0 \left[ 1 - \frac{S_\infty}{N} \right], \quad (1.7)$$

kde

$$\ln \left( \frac{S_0}{S_\infty} \right) = \mathcal{R}_0 \left[ 1 - \frac{S_\infty}{N} \right]$$

se nazývá vztah konečné velikosti (*final size relation*), neboť dává přímý vztah mezi základním reprodukčním číslem a rozsahem epidemie. Závorka pravé strany rovnice

bývá také někdy označována jako míra zasažení (*attack rate*), jenž je vyjádřením poměru jaká část populace byla epidemií postižena [6].

Při změně mezí integrace na konečný čas dospějeme stejným postupem jako výše ke

$$I(t) + S(t) - \frac{\alpha}{\beta} \ln(S(t)) = N - \frac{\alpha}{\beta} \ln(S_0). \quad (1.8)$$

Jedná se o implicitní funkci, kterou můžeme zobrazit graficky pro různé počáteční podmínky ve fázové  $SI$  rovině (viz obrázek 1.1). Jestliže na počátku epidemie platí  $\mathcal{R}_0 > 1$  (stejně jako  $I'(0) > 0$ ), derivací a nalezením extrémů ( $S = \alpha/\beta$ ) pro rovnici (1.8), obdržíme maximální velikost kompartmentu infikovaných během epidemie

$$I_{max} = S_0 + I_0 - \frac{\alpha}{\beta} \ln(S_0) - \frac{\alpha}{\beta} + \frac{\alpha}{\beta} \ln\left(\frac{\alpha}{\beta}\right). \quad (1.9)$$

Zřejmě platí závislost koeficientu  $\alpha$  na charakteru choroby, neboť určuje průměrnou dobu setrvání v kompartmentu nemocných, tedy lze očekávat určení ze sérologického výzkumu. Jak bylo řečeno,  $\beta$  parametr považujeme zatím za konstantu, avšak spíše by bylo přirozené uvažovat závislost  $\beta(N)$  se saturací pro velká  $N$  a dále  $\beta'(N) < 0$  od určitého  $N_0$ . Autoři [9] statisticky vyhodnocují některé závislosti  $\beta$ , jako jsou

$$\beta(N) = \frac{a}{1 + bN}, \quad \text{či} \quad \beta(N) = aN^{-0.95}. \quad (1.10)$$

První vztah je analogií k interakcím typu Michaelise-Mentenové, kde  $a$  je standardní incidence, respektive  $b$  konstanta úměrnosti. Druhý výraz autoři dovodili na základě dat z velkých měst a postulují toto chování pro všechny města od určité velikosti. Zřejmě lze dále čekat promítnutí protiepidemických opatřeními, biologické specifity a behaviorálních faktorů do koeficientu  $\beta$ . U mnoha chorob je nutné uvažovat i oscilační charakter  $\beta$  (například u dětských infekcí návrat dětí do škol apod.).

Sérologické testy mohou posloužit k určení  $S_0$  a  $S_\infty$  před a po pandemii. Poté je možné retrospektivně odhadnout celkové  $\mathcal{R}_0$  podle (1.7). Pokud uvažujeme funkci (1.7) s proměnou  $S_\infty$ , můžeme dospět k nerovnostem, které mohou být použity již během epidemie (při odhadu  $\mathcal{R}_0$  z aktuálních dat - viz dále v textu) [4].

$$S_\infty < \frac{S_0}{\mathcal{R}_0}, \quad \text{resp.} \quad S_\infty < \frac{N}{\mathcal{R}_0}. \quad (1.11)$$

Z analytického hlediska je možné dokázat neexistenci limitních cyklů pro systém 1.4, i když připustíme například návrat jedinců do  $S$  úměrný  $R$  (dejme  $\chi R$ ) či proměnou  $\beta(t)$ . Například při použití Dulacova kritéria  $\vec{\nabla}g(x,y)d\vec{x}/dt$ , kde  $g$  je vhodně zvolená, spojitá funkce na nějaké jednoduše spojitě oblasti  $\Omega$ , resp.  $d\vec{x}/dt$  je spojitý systém. Jestliže daná divergence funkce výše nemění na oblasti  $\Omega$  znaménko, potom neexistuje uzavřený orbit v  $\Omega$  (důkaz například v [7]). Dostáváme pro oblast

$S > 0, I > 0$ , systém (1.4) a funkci  $g = 1/I$

$$\frac{\partial}{\partial S} \left( -\beta(t)S + \frac{\chi R}{I} \right) + \frac{\partial}{\partial I} (\beta(t)S - \alpha) = -\beta(t) < 0, \quad (1.12)$$

neboť frekvence kontaktů ( $\beta(t)$ ) nemůže být záporná.

SIR model můžeme dále zobecňovat postulováním zobecněné pravděpodobnostní distribuce  $P(s)$ , jenž vyjadřuje jaká část nemocných v čase  $s$  po nakažení je stále infekční pro třídu náchylných. Jinými slovy zavádíme pravděpodobnostní rozdělení, které určuje šanci setrvání v kompartmentu nemocných v čase  $s$  po nakažení, tj. neomezujeme se na exponenciální rozdělení jako doposud (šířeji pojednáno v sekci 1.5). Rovnice (1.4) po této úpravě dává

$$\begin{aligned} S' &= -\beta SI \\ I' &= I_0 + \int_0^t -S'(t-s)P(s) ds. \end{aligned} \quad (1.13)$$

(Protože  $-S'(t-s)$  je počet nových infekcí v čase  $(t-s)$ .) Nechť je  $P(s)$  nezáporná, nerostoucí funkce s  $P(0) = 1$  a její průměrná hodnota, tj.  $\int_0^\infty P(s) ds$  (či jinak  $\int_0^\infty -sP'(s) ds$ ), je konečná. Protože se jedná o analogické vyjádření  $1/\alpha$  (průměrná doba, kterou je jedinec nemocný) ze (1.4), můžeme vyjádřit nově reprodukční číslo jako

$$\mathcal{R}_0 = \beta N \int_0^\infty P(s) ds. \quad (1.14)$$

Další možností, avšak podobnou, je ta popsaná Kermackem a McKendrickem [3]. Uvažovali znovu o infektivitě, která je ovšem závislá na době, po kterou je jedinec nemocný (*age of infection*). Nechť  $A(s)$  značí část třídy infikovaných, jenž jsou nakažlivý v čase  $s$  po nakažení a odráží také pravděpodobnost, že kontakt povede k vytvoření nově infikovaného jedince. Zavedme  $\varphi(t)$  vyjadřující celkovou infektivitu, tj. počet infikovaných a jejich schopnost infikovat

$$\varphi(t) = \int_0^t -S'(t-s)A(s) ds. \quad (1.15)$$

(předpokládáme první nakažené až skokově v čase  $t = 0$ ). SIR můžeme poté přepsat na

$$\begin{aligned} S'(t) &= -\beta S(t)\varphi(t) \\ \varphi(t) &= \int_0^t -S'(t-s)A(s) ds. \\ \Rightarrow S'(t) &= \beta S(t) \int_0^t S'(t-s)A(s) ds. \end{aligned} \quad (1.16)$$

Jak je tedy patrné,  $\mathcal{R}_0$  pro (1.16) činí  $\beta N \int_0^\infty A(s) ds$ . Pokusíme se odhadnout  $\mathcal{R}_0$  na počátku epidemie, kdy dochází k exponenciálnímu růstu  $I(t)$ . Nechť počáteční pokles  $S(t)$  nesnižuje frekvenci kontaktů s infikovanými (tedy  $\beta S(t) \approx \beta N$ ). Kladme



na poslední rovnici (1.16) podmínku, abychom dostali řešení ve tvaru  $S(t) = Ne^{rt}$ , což dává

$$\beta N \int_0^\infty A(s) e^{-rs} ds = 1, \quad (1.17)$$

kde  $r$  udává rychlost poklesu (možno odhadnout fitem dat). Protože sérologická vyšetření mohou pro danou chorobu sloužit jako odhad pro  $A(s)$ , dostáváme vydělením  $\mathcal{R}_0$  a (1.17) vztah

$$\mathcal{R}_0 = \frac{\int_0^\infty A(s) ds}{\int_0^\infty A(s) e^{-rs} ds}. \quad (1.18)$$

Vidíme  $\mathcal{R}_0 > 1$  tehdy a jen tehdy platí-li  $r > 0$ .

## 1.3 SEIR model

Model SIR je možné zpřesňovat přidáváním dalších kompartmentů společně s definicí toku jedinců mezi nimi, jenž je řízen novými koeficienty, resp. velikostmi zbývajících kompartmentů v čase. Jedna z možností je rozšíření o nemocné jedince, kteří ovšem ještě nejsou infekční, tzv. *exposed* (značme  $E$ ). Necht je průměrná doba, jenž stráví infikovaní v této třídě  $1/\kappa$  [4]. Poté dostáváme systém

$$\begin{aligned} S' &= -\beta SI \\ E' &= \beta SI - \kappa E \\ I' &= \kappa E - \alpha I \end{aligned} \quad (1.19)$$

(Znovu vynecháváme třídu vyléčených, neboť je velikost populace konstantní). Tento model již není možné plně analyticky vyřešit. Můžeme se však pokusit o *linearizaci* okolo kritických bodů (ekvilibríí) systému [14]. Pokud provedeme Taylorův rozvoj systému okolo kritických bodů (body ve fázovém prostoru, pro něž platí  $(S', E', I') = (0, 0, 0)$ ) a předpokladem spojitosti a diferencovatelnosti funkce systému v tomto okolí, potom může být systém (1.19) velmi dobře aproximován lineárním diferenciálním systémem, jenž je vyjádřen *Jakobiánem* ( $\mathbf{J}$ ) v tomto kritickém bodě. Tedy touto aproximací se dopustíme maximálně chyby řádu  $\mathcal{O}(\eta^2)$ , kde  $\eta$  představuje vzdálenost od kritického bodu. Tato chyba je velmi malá, neboť předpokládáme malou odchylku od tohoto bodu.

Každou reálnou matici  $\mathbf{A}$  s různými vlastními čísly je možné diagonalizovat vůči určité bázi jejího vektorového prostoru, tedy vyjádřit jako  $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1}$ , kde operátor  $\mathbf{V}$  je tvořen vlastními vektory  $\mathbf{A}$ , respektive  $\mathbf{\Lambda}$  je diagonální matice obsahující vlastní čísla  $\mathbf{A}$  [13]. Formálně můžeme tedy diferenciální systém prvního řádu i s řešením zapsat jako

$$\frac{d\vec{x}(t)}{dt} = \mathbf{A}\vec{x}(t) \quad \Rightarrow \quad \vec{x}(t) = \vec{x}(0)e^{\mathbf{A}t} \quad \Rightarrow \quad \vec{x}(t) = \vec{x}(0)\mathbf{V}e^{\mathbf{\Lambda}t}\mathbf{V}^{-1} \quad (1.20)$$

Systém (1.19) (označme  $\mathbf{S}$ ) má pouze jedno ekvilibrium a to pro  $S = N, E = 0, I = 0$ , které vyjadřuje vymýcení choroby. Linearizace dává

$$\begin{aligned} \begin{bmatrix} S \\ E \\ I \end{bmatrix}' &\approx \mathbf{J}_{\mathbf{S}}(N, 0, 0) \begin{bmatrix} S \\ E \\ I \end{bmatrix} = \begin{bmatrix} \frac{\partial S'}{\partial S} & \frac{\partial S'}{\partial E} & \frac{\partial S'}{\partial I} \\ \frac{\partial E'}{\partial S} & \frac{\partial E'}{\partial E} & \frac{\partial E'}{\partial I} \\ \frac{\partial I'}{\partial S} & \frac{\partial I'}{\partial E} & \frac{\partial I'}{\partial I} \end{bmatrix} (N, 0, 0) \begin{bmatrix} S \\ E \\ I \end{bmatrix} = \\ &= \begin{bmatrix} 0 & 0 & -\beta N \\ 0 & -\kappa & \beta N \\ 0 & \kappa & -\alpha \end{bmatrix} \begin{bmatrix} S \\ E \\ I \end{bmatrix} \end{aligned} \quad (1.21)$$

Vlastní čísla určíme nalezením kořenů charakteristického polynomu matice výše

$$\begin{aligned} -\lambda(\kappa + \lambda)(\alpha + \lambda) + \lambda\kappa\beta N &= -\lambda(\lambda^2 + \lambda(\alpha + \kappa) - \kappa(\beta N - \alpha)) = 0 \\ \Rightarrow \lambda_1 &= 0 \quad ; \quad \lambda_{2,3} = \frac{-(\alpha + \kappa) \pm \sqrt{(\alpha - \kappa)^2 + 4\kappa\beta N}}{2}. \end{aligned} \quad (1.22)$$

Cílem celého snažení je určit, zda okolo ekvilibria je řešení asymptoticky stabilní (zánik epidemie) či nestabilní (propuknutí epidemie). To určují vlastní čísla dovozená výše. Je zřejmé, že stabilita nastává v tomto případě pouze, pokud je stopa Jakobiánu záporná, resp. determinant rohové nenulové  $2 \times 2$  matice ( $\kappa(\alpha - \beta N)$ ) Jakobiánu je kladný. (Neboť negativní reálné části vlastních čísel mají za následek konvergenci k ekvilibriu systému (tzv. *sink*, popřípadě *spiral sink*, jestliže má vlastní číslo imaginární část). Lze dokázat identické chování systému s jeho linearizovanou formou okolo kritických bodů v případech zmíněných a dále pro *saddle*, *source* a *spiral source*. Avšak například u čistě imaginárních vlastních čísel (*centers*) je nutné zvolit jiný postup jako je využití *Poincarého-Bendixovy věty*, protože se chování mohou lišit. V těchto bodech vyvstávají například *Hopfovy bifurkace* [14, 13, 7].

Všimněme si, že zmíněný determinant výše je kladný pouze když  $\mathcal{R}_0 < 1$  (jinak řečeno  $\kappa\alpha(1 - \mathcal{R}_0)$ ), kde  $\mathcal{R}_0$  je stejně formulováno jako v původním SIR modelu. V opačném případě ( $\mathcal{R}_0 > 1$ ) můžeme tvrdit, že zavedení kompartmentu  $E$  má za následek zpomalení počátečního exponenciálního růstu infekce, jak je patrné z vyjádření  $\lambda_2$  výše.

Podobně jako u SIR modelu je i zde možné zavést různá zobecnění [6]. Necht funkce  $P_E(t)$ , respektive  $P_I(t)$ , udávají pravděpodobnost přetrvání ve třídě  $E$ , respektive  $I$ , tedy ať jsou nezáporné, nerostoucí a dále

$$P_E(0) = P_I(0) = 1; \quad \int_0^\infty P_E(s) ds < \infty; \quad \int_0^\infty P_I(s) ds < \infty. \quad (1.23)$$

O záporně vzatých derivacích těchto funkcí můžeme uvažovat jako o hustotách pravděpodobnosti pro jednotlivé kompartmenty. Nové vyjádření  $E(t)$  je dáno jako

$$E(t) = E_0 P_E(t) + \int_0^t -S'(t - u) P_E(u) du. \quad (1.24)$$

Derivací rovnice výše a společně se zjednodušujícím předpokladem, že došlo v čase  $t = 0$  ke skokové změně  $S'(t)$  o  $-E_0$ , dostáváme

$$E'(t) = \int_0^t -S'(t-u)P'_E(u) du, \quad (1.25)$$

kde bylo využito faktu o derivaci konvoluce dvou funkcí  $((-S' * P_E)' = -S' * P'_E)$  a komutativity konvoluce [7]. Tento výraz ovšem také vyjadřuje vstup do infekční fáze z  $E$  skupiny. Potom  $I(t)$  udává rovnice

$$\begin{aligned} I(t) &= - \int_0^t \int_0^u -S'(s)P'_E(u-s)P_I(t-u) ds du = \\ &= - \int_0^t \int_s^t -S'(s)P'_E(u-s)P_I(t-u) du ds \\ \Rightarrow I(t) &= \int_0^t -S'(s)A(t-s) ds, \end{aligned} \quad (1.26)$$

kde

$$A(t-s) = - \int_s^t P'_E(u-s)P_I(t-u) du = \int_0^{t-s} P'_E(t-s-v)P_I(v) dv. \quad (1.27)$$

Model SEIR dostává následující podobu

$$\begin{aligned} S' &= \beta SI \\ E(t) &= \int_0^t -S'(t-s)P_E(s) ds \\ I(t) &= \int_0^t -S'(t-s)A(s) ds. \end{aligned} \quad (1.28)$$

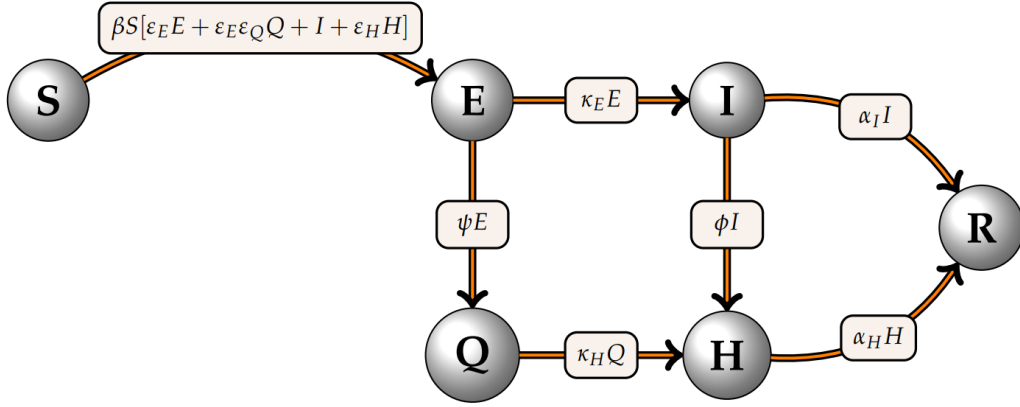
Reprodukční číslo poté činní

$$\begin{aligned} \mathcal{R}_0 &= \beta N \int_0^\infty A(z) dz = -\beta N \int_0^\infty \int_0^z P'_E(z-u)P_I(u) du dz = \\ &= \beta N \int_0^\infty P_I(u) du, \end{aligned} \quad (1.29)$$

protože platí  $\int_0^\infty -P'_E(u) du = P_E(0) - P_E(\infty) = 1$ .

## 1.4 Modelování karantény a izolace

V případě propuknutí epidemie nové nemoci či nemoci způsobené známým avšak zmutovaným patogenem, je jedinou možností zastavení šíření (vakcinace ani žádná účinná léčba není dostupná) izolace nakažených jedinců a karanténa těch, jenž s nimi mohli přijít do kontaktu. Za účelem předpovědi chování epidemie SARS v Číně roku 2002-2003 byl představen model, který kvantifikuje tyto opatření. Jedná se o rozšířený model SEIR. Necht  $E$  je třída nově infikovaných  $E$  (*exposed*) infekční avšak redukovanou faktorem  $\varepsilon_E$ , a přestupují do třídy infikovaných  $I$  rychlostí  $\kappa_E$ ,



Obr. 1.2: Diagram modelu SEQIHR

pokud nejsou v izolaci. Třída jedinců v karanténě  $Q$  je složena z členů populace, jenž byli testováni či se setkali s nakaženým. Karanténa není zcela účinná, avšak předpokladem je snížení počtu kontaktů faktorem  $\varepsilon_Q$ . Obecně jsou samozřejmě často poslání do karantény i zdravý jedinci, ale pro zjednodušení vynechejme. Necht jsou členové kompartmentu  $I$  diagnostikováni rychlostí  $\phi$  za jednotku času a poslání do třídy izolovaných  $H$  (v literatuře také *hospitalized*). Podobně při rozvinutí příznaků u jedinců v karanténě, jenž jsou monitorováni, bude jejich přechod do třídy  $H$  řízen faktorem  $\kappa_Q$ . Necht je infekčnost kompartmentu  $H$  snížena faktorem  $\varepsilon_H$ . Infikování, respektive izolování jsou uzdravování rychlostí  $\alpha_I$ , respektive  $\alpha_H$ . Tedy můžeme formulovat systém

$$\begin{aligned}
 S' &= -\beta S[\varepsilon_E E + Q + I + \varepsilon_H H] \\
 E' &= \beta S[\varepsilon_E E + \varepsilon_E \varepsilon_Q Q + I + \varepsilon_H H] - (\kappa_E + \psi)E \\
 Q' &= \psi E - \kappa_Q Q \\
 I' &= \kappa_E E - (\alpha_I + \phi)I \\
 H' &= \kappa_Q Q + \phi I - \alpha_H H
 \end{aligned} \tag{1.30}$$

(podle [6]). V případech komplikovanějších modelů je výhodné zobrazovat tok jedinců mezi kompartmenty diagramy (pro (1.30) na Obr. 1.2).

Například při linearizaci okolo kritického bodu, jenž vyjadřuje populaci prostou infekce, dostáváme Jakobián v tomto bodě

$$\begin{bmatrix}
 \varepsilon_E \beta N - (\kappa_E + \psi) & \varepsilon_E \varepsilon_Q \beta N & \beta N & \varepsilon_H \beta N \\
 \psi & -\kappa_Q & 0 & 0 \\
 \kappa_E & 0 & -(\alpha_I + \phi) & 0 \\
 0 & \kappa_Q & \phi & -\alpha_H
 \end{bmatrix}, \tag{1.31}$$

kde byl vynechán první sloupec, respektive první řádek, neboť byl nulový, respektive s jedním nenulovým prvkem, který neležel na diagonále, takže kořeny charakteristického polynomu zůstávají nezměněny (samozřejmě opomínáme kořen  $\lambda = 0$ , který

zůstává touto úpravou nevyjádřen). Definujeme reprodukční číslo, které bude odrážet až začátek rozpoznání epidemie a přijetí protiepidemických opatření, značme  $\mathcal{R}_c$  (*control reproduction number*). Jinak řečeno, předpokládáme faktory, které řídí skupiny populace v karanténě a izolaci, za rovny nule, dokud nebyly zavedeny během epidemie. Pokud bychom našli všechny čtyři kořeny charakteristického polynomu matice výše, dospěli bychom znovu k existenci čísla, které řídí, zda všechny vlastní čísla mají negativní reálné části. Explicitně vyjádřeno

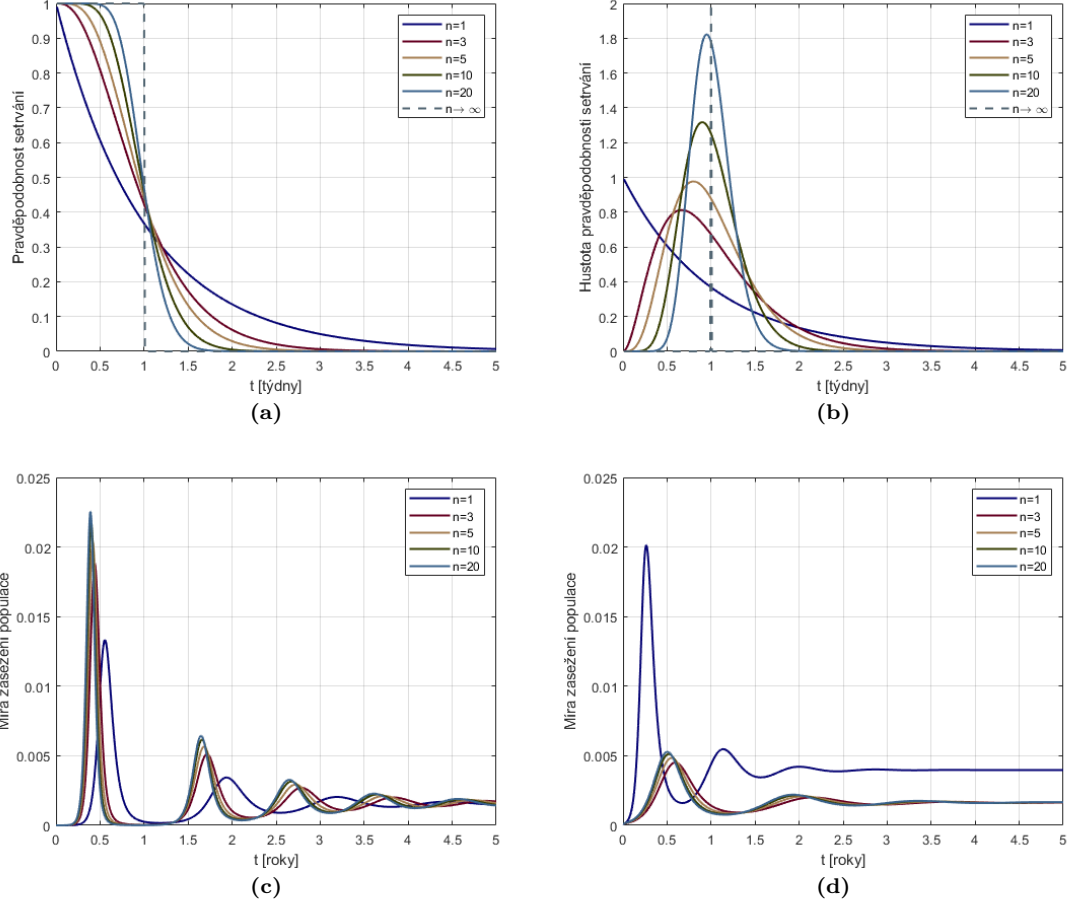
$$\mathcal{R}_c = \beta N \left[ \frac{\varepsilon_E}{\psi + \kappa_E} + \frac{\kappa_E}{(\psi + \kappa_E)(\phi + \alpha_I)} + \frac{\varepsilon_Q \varepsilon_E \psi}{\kappa_Q(\psi + \kappa_E)} + \frac{\varepsilon_H \kappa_E \phi}{\alpha_H(\psi + \kappa_E)(\phi + \alpha_I)} + \frac{\varepsilon_H \psi}{\alpha_H(\psi + \kappa_E)} \right] \quad (1.32)$$

(podrobné dovození v [4]). Jak je patrné z rovnic (1.30), parametry  $\psi$  a  $\phi$  jsou klíčové k řízení protiepidemických opatření. Faktory  $\varepsilon_Q$ , respektive  $\varepsilon_H$  jsou vyjádřením efektivnosti daných opatření. Další parametry jsou specifické pro danou chorobu. Každému zlomku ve vyjádření  $\mathcal{R}_c$  můžeme přiřadit epidemiologickou interpretaci. Tedy znovu se jedná o sekundární infekce, které by způsobil jeden nakažený v plně náchylné populaci. Můžeme uvažovat následovně. Každý zlomek v  $\mathcal{R}_c$  má v čitateli průměrný počet kontaktů v jednotlivých kompartmentech nebo jejich kombinacích (jen ty kombinace, které jsou možné při toku jedinců mezi nimi). Podobně jmenovatele jsou mírou jednotlivých průměrných časů strávených v těchto kompartmentech či jejich kombinacích. Tedy jmenovitě první zlomek udává příspěvek jedinců z  $E$ , druhý těch putujících z  $E$  do  $I$ , třetí z  $E$  do  $Q$ , čtvrtý z  $E$  přes  $I$  do  $H$  a konečně pátý z  $E$  přes  $Q$  do  $H$ .

## 1.5 Gama distribuce infekčnosti

Mnoho studií poukazuje na biologickou nerealističnost exponenciálního rozložení času stráveného jedincem v daném kompartmentu pro většinu infekčních chorob [10]. Jak bylo dokázáno v předchozích sekcích, je možné lehce přeformulovat model pro obecnou distribuci  $P(t)$ . (Lze dokázat existenci globálního atraktoru pro  $\mathcal{R}_c < 1$ , resp.  $\mathcal{R}_0 < 1$ , jenž vyjadřuje populaci prostou infekce, i pro systému s obecnou distribucí. Vice versa pro  $\mathcal{R}_c > 1$ , resp.  $\mathcal{R}_0 > 1$  existuje atraktor představující endemické ekvilibrium [10].) Uvažujme například rozdělení infekčnosti jedinců na  $n$  částí s průměrnou dobou setrvání  $1/n\alpha$ . Nechť  $u_k(\tau)$  představuje zlomek jedinců, jenž jsou infekční po dobu  $\tau$  v  $k$ -tém subintervalu. Potom zřejmě musí platit

$$\begin{aligned} u_1' &= -n\alpha u_1, \quad \text{pro } u_1(0) = 1 \\ u_k' &= n\alpha(u_{k-1} - u_k), \quad \text{pro } u_1(0) = 0, \quad k = 2, \dots, n. \end{aligned} \quad (1.33)$$



**Obr. 1.3:** (a)+(b) vyobrazení pravděpodobností, respektive jejich hustoty setrvání jedince v daném kompartmentu pro různé  $n$  a  $1/\alpha = 1$ . (c) porovnání simulace SEIR modelu pro exponenciální ( $m = n = 1$ ) a gama distribuce ( $m = 5; n > 1$ ) pro  $N = 10^6$ ,  $E(0) = R(0) = 0$ ,  $I(0) = 1$ ,  $1/\kappa = 2$  dny,  $1/\alpha = 10$  dní,  $\mu = 0.0005$ .  $\beta$  zvoleno, aby platilo  $\mathcal{R}_0 = 2$ . (d) ta samá simulace pro  $\mathcal{R}_0 = 1.7$ . Zasažení vyjádřeno jako počet infikovaných vůči celé populaci.

Vyřešením rekurzivně ihned dostáváme kernel

$$P(t) = \sum_{k=1}^n u_k(t) = \sum_{k=1}^n \frac{(n\alpha)^{k-1}}{\Gamma(k)} t^{k-1} e^{-n\alpha t}, \quad (1.34)$$

kde  $\Gamma(z)$  je gama funkce. Tedy (1.34) vyjadřuje gama distribuci s tvarovým parametrem  $n$  a škálovým parametrem  $1/\alpha$ . Je očividný krajní případ pro  $n = 1$ , kdy vztah přechází do exponenciálního rozložení. Naopak pro  $n \rightarrow \infty$  se doba setrvání blíží konstantní hodnotě  $1/\alpha$ . Pokud předpokládáme takto specifickou distribuci jako gama, potom je možné obecně formulované systémy jako (1.28) převést zpět do diferenciální formy.

Ukázka gama rozdělení pro vybraná  $n$ , respektive jejich hustoty jsou vyobrazeny na Obr. 1.3a, respektive Obr. 1.3b. Dále jsou na Obr. 1.3c a Obr. 1.3d některá srovnání numerických simulací modelu SEIR 1.19 ve formulaci pro gama distribuci.

To znamená rozčlenění kompartmentu  $E$  na  $m$  fází a kompartment  $I$  na  $n$  fází v duchu rovnic (1.33). (Bylo například zjištěno  $m = n = 20$  pro spalničky (pro  $1/\kappa = 8$  dní,  $1/\alpha = 5$  dní), či  $m = 40; n = 4$  pro pravé neštovice (pro  $1/\kappa = 14$  dní,  $1/\alpha = 8.6$  dne) [10].) Všimněme si krajního případu  $m = 1; n = 1$ , kdy rozdělení přechází v exponenciální. Aby byly diskrepance mezi těmito případy více zdůrazněny, bylo umožněn tok jedinců do  $S$  úměrný celkové velikosti populace ( $\mu N$ ), respektive tok z kompartmentů úměrný jejich velikosti škálován také parametrem  $\mu$ . Dynamika modelu poté získává pro  $\mathcal{R}_0 > 1$  podobu tlumených oscilací. U obou SEIR simulací jsou jasně patrné odlišnosti mezi exponenciálním a gama rozdělením. V prvním případě byl do populace představen 1 infekční jedinec s předpokladem  $\mathcal{R}_0 = 2$ . Vidíme populačně extenzivnější avšak kratší proběhnutí první vlny epidemie v případě gama a opožďování vln s exponenciálním rozdělením. Ve druhém případě došlo ke snížení reprodukčního čísla na  $\mathcal{R}_0 = 1.7$ . Vidíme znovu signifikantně odlišné chování pro obě rozdělení, nejen co v míře zasažení pro jednotlivé vlny, zpoždění v nástupu vln, ale i v různosti endemických ekvilibrií. (podrobnější srovnání v sekci 1.7)

## 1.6 Počátek epidemie a *network based* modely

Nedostatkem deterministické kompartmentové formulace je zejména počátek epidemie, neboť se šíření týká jen velmi malého množství jedinců a deterministický model předpokládá, že každý jedinec populace má stejnou pravděpodobnost kontaktu s již nakaženými jedinci, což v případě malé skupiny infikovaných neodpovídá realitě. Aby bylo možné tuto skutečnost zahrnout, je nutné vyhodnotit demografickou stochasticitu, tj. předtím než dojde k přibližnému exponenciálnímu růstu počtu nakažených, vytvořit model, který tuto proměnou bude zahrnovat (celá sekce z [8, 4, 11, 12, 6]).

Budeme uvažovat o členech skupiny jako vrcholech grafu (neorientovaného - připouštíme šíření oběma směry), kde hrany mezi nimi budou představovat jednotlivé kontakty. Použijeme postup velmi podobný *Galtonovu-Watsonovu procesu*, který se snažil modelovat pravděpodobnost extinkce aristokratických příjmení patrilineárně ve viktoriánské Anglii. My se podobně budeme ptát na pravděpodobnost zastavení šíření infekce po grafu jedinců. (Taková abstrakce je obecně velmi vhodná i pro popis šíření sexuálně přenosných chorob, kde se ukazuje, že v síti kontaktů působí jen několik velmi aktivních přenašečů a u ostatních jedinců existuje pouze jediné spojení a to s těmito přenašeči).

Nechť  $p_k$  představuje pravděpodobnost, že náhodně vybraný jedinec bude mít právě  $k$  kontaktů (stupeň vrcholu) s dalšími členy skupiny. Potom označme  $p_k$  pravděpodobnostní rozdělení pro všechny množství styků při náhodném výběru jedince.

Protože se jedná o diskretní náhodnou proměnou, je pohodlnější pracovat s její vytvořující funkcí (*probability-generating function*)[12]. Tedy označme

$$G_0(z) = \sum_{k=0}^{\infty} p_k z^k, \quad (1.35)$$

kde  $0 \leq z \leq 1$  a zřejmě platí  $\sum_{k=0}^{\infty} p_k = 1$ . Dále definujme průměrný stupeň (první moment) vrcholu jako

$$\langle k \rangle = \sum_{k=0}^{\infty} k p_k = G'_0(1). \quad (1.36)$$

Předpokládáme, že infekce byla přenesena a začala v jednom vrcholu grafu (*patient zero*). Protože jedinec nemůže šířit chorobu po hraně, která ho infikovala, zavedme přebytečný stupeň (*excess degree*) daného vrcholu grafu. Pokud má vrchol stupeň  $k$ , tak jeho přebytečný stupeň činí  $(k - 1)$ . Jestliže náhodně vybereme hranu grafu, potom je pravděpodobnost, že jeden z vrcholů, který spojuje bude mít přebytečný stupeň  $(k - 1)$  je konstanta krát  $k p_k$ . Necht' je tato konstanta  $1/\langle k \rangle$ , aby tento výběr tvořil pravděpodobnostní rozdělení. Tedy vyjádříme tuto pravděpodobnost jako

$$q_{k-1} = \frac{k p_k}{\langle k \rangle} \quad (1.37)$$

a její vytvořující funkci jako

$$G_1(z) = \sum_{k=1}^{\infty} q_{k-1} z^{k-1} = \sum_{k=1}^{\infty} \frac{k p_k}{\langle k \rangle} z^{k-1} = \frac{1}{\langle k \rangle} G'_0(z). \quad (1.38)$$

Velkou důležitost bude v další úvaze hrát průměrný přebytečný stupeň

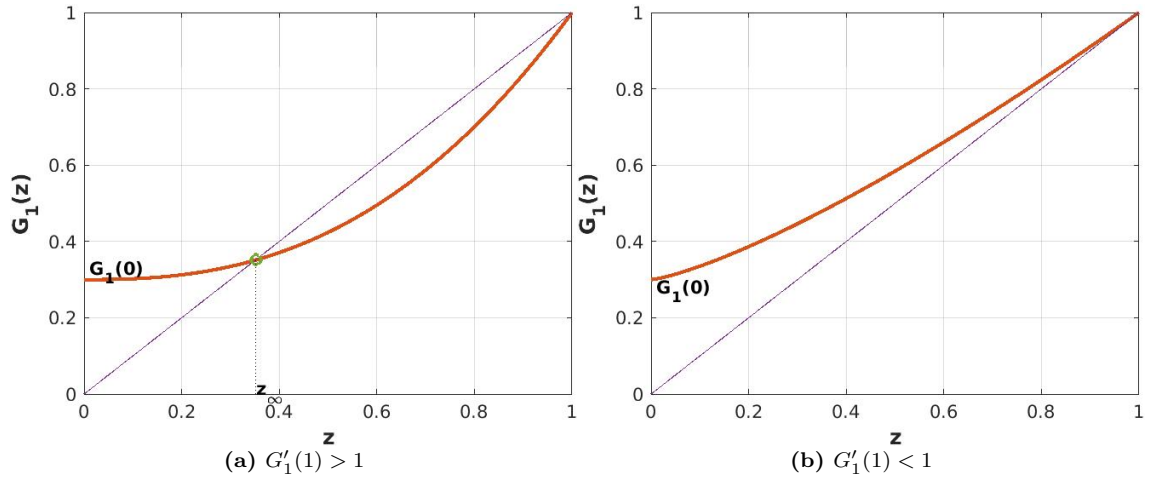
$$\begin{aligned} \langle k_e \rangle &= \frac{1}{\langle k \rangle} \sum_{k=1}^{\infty} k(k-1) p_k = \\ &= \frac{1}{\langle k \rangle} \sum_{k=1}^{\infty} k^2 p_k - \frac{1}{\langle k \rangle} \sum_{k=1}^{\infty} k p_k = G'_1(1). \end{aligned} \quad (1.39)$$

Konečně necht' sekundárně nakažený vrchol má přebytečný stupeň  $j$ . Potom necht'  $z_n$  značí pravděpodobnost, že infekce zanikne v populaci po  $n$  přenosech (po  $n$ -krokové procházce po grafu). Tj. ovšem stejné jako tvrzení, že všechny sekundární infekce (ty pocházející od nulového vrcholu), musí zaniknout v  $(n - 1)$  přenosech. Dostáváme tedy rekurentní vztah

$$z_n = \sum_{j=0}^{\infty} q_j z_{n-1}^j = G_1(z_{n-1}). \quad (1.40)$$

Předpokládejme, že došlo k přenosu, tedy  $z_0 = 0$ . Dále necht'  $G_1(0) > 1$ , což implikuje neexistenci vrcholů grafu pouze s jednou hranou. Všimněme si platnosti  $G'_1(z) > 0$  a  $G''_1(z) > 0$  pro všechna  $z$ . Tedy  $z_n$  je rostoucí posloupnost a funkce  $G_1(z)$  je konkávní na intervalu  $[0, 1]$ . Protože  $z_n$  je rostoucí a omezená musí mít





**Obr. 1.4:** Grafická interpretace existence kořenů  $z = G_1(z)$ .

limitu  $z_\infty$ .  $z_\infty$  musí být ekvilibriem diferenční rovnice (1.11) jinými slovy řešením  $z = G_1(z)$ . Ihned je jasné jedno řešení  $z = 1$ . Avšak platí-li  $G'_1(1) > 1$  musí existovat řešení  $z < 1$ , neboť  $(z)' = 1 > G'_1(1)$  společně s  $z = 0 < G_1(0)$  a konkávností  $G_1(z)$  to podmiňují (viz Obrázek 1.4).

Můžeme tedy připsat roli  $\langle k_e \rangle$  jako číslu  $\mathcal{R}_0$  v kompartmentovém modelu. Rozdílnost je však pro  $\langle k_e \rangle > 1$ , kdy existuje nenulová pravděpodobnost, že dojde k zastavení epidemie, což není možné u kompartmentového modelu ( $\mathcal{R}_0 > 1$ ). (Pokud bychom pouze chtěli kategoricky rozhodnout o téměř jistém vzniku epidemie a vice versa, můžeme uvažovat jednodušeji. Necht  $X_n$  značí počet případů infekce po  $n$  krocích po grafu a  $Y_{n-1,i}$  je počet nově nakažených, který vznikl od jedince  $i$  v předchozím kroku. Pro očekávanou hodnotu  $X_n$  za předpokladu, že bylo nakaženo právě  $k$  osob v předchozím kroku, platí

$$\mathbf{E}[X_n | X_{n-1} = k] = \mathbf{E} \left[ \sum_{i=1}^k Y_{n-1,i} | X_{n-1} = k \right] = k \mathbf{E}[Y] = k\mu, \quad (1.41)$$

neboť počet nově nakažených (počet hran), není závislý na předchozím počtu infikovaných. Můžeme tedy formulovat obecnou očekávanou hodnotu pro  $X_n$

$$\begin{aligned} \mathbf{E}[X_n] &= \mathbf{E}[\mathbf{E}[X_n | X_{n-1}]] = \sum_{k=0}^{\infty} \mathbf{P}[X_{n-1} = k] \mathbf{E}[X_n | X_{n-1} = k] = \\ &= \mu \sum_{k=0}^{\infty} k \mathbf{P}[X_{n-1} = k] = \mu \mathbf{E}[X_{n-1}] = \mu^n \mathbf{E}[X_0] = \mu^n, \end{aligned} \quad (1.42)$$

kde je v prvním kroku použito *Adamova zákona*. Protože počátek epidemie začíná jediným člověkem, platí  $X_0 = 1$ . Tedy jestliže platí  $\mu < 1$ , epidemie nevznikne. Je zřejmá ekvivalence s definicí  $\langle k_e \rangle$ .

Pro ilustraci uvažujme například situaci, kdy 6 z 10 jedinců neinfikují žádného jedince, 1 z deseti jednoho jedince, resp. 5 jedinců, resp. 7 jedinců. Potom platí

$$G_0(z) = \frac{7}{10} + \frac{z}{10} + \frac{z^5}{10} + \frac{z^7}{10}; \quad G_1(z) = \frac{1}{13} + \frac{5z^4}{13} + \frac{7z^6}{13}; \quad z_\infty = 0.077. \quad (1.43)$$

Zpětným dosazením do  $G_0$  (vyjadřuje pravděpodobnost, že infekce přenesena do populace jediným jedincem, nezpůsobila epidemii) dává  $G_0(z_\infty) \approx 0.7$ . Pravděpodobnost vypuknutí epidemie je poté přibližně 30%.

Pokud bychom vybrali náhodnou skupinu osob a určily pacienta nula, od kterého by se nemoc šířila, je nejvhodnější uvažovat o grafu kontaktů jako o *grafu Erdős – Rényi*. Potom kontakt dvou osob je přibližně náhodný a při velkém počtu vrcholů má vrchol pravděpodobnost  $k$  hran limitně *Poissonovo rozdělení*, konkrétně

$$p_k = \frac{e^{-\langle k_e \rangle} \langle k_e \rangle^k}{k!}. \quad (1.44)$$

Tento fakt zjednodušuje výpočty, neboť jednoduchá manipulace ukáže, že platí  $G_0(z) = G_1(z)$ , resp.  $\langle k \rangle = \langle k_e \rangle$  a navíc

$$G_0(z) = e^{\mathcal{R}_0(z-1)}, \quad (1.45)$$

kde pro vyhnutí se ambivalenci značme  $\langle k \rangle = \mathcal{R}_0$ . Při aplikaci na příklad výše bychom dostali  $z_\infty \approx 0.577$ , poté by pravděpodobnost epidemie činila přibližně 42%. Vidíme, že se výsledky velmi odlišují. To je způsobeno existencí superpřenašečů v prvním případě.

Doteď jsme předpokládali absolutní efektivnost, pokud dojde ke kontaktu s nakaženým. Zavedme číslo  $T$ , jenž bude značit *transmisibilitu* (dosud  $T = 1$ ). Biologicky by mělo odrážet čas, po který je jedinec nakažený, náchylnost jedince k infekci, rychlost jakou kontakty probíhají apod. Nechť jsou hrany, které přenesly chorobu mezi vrcholy tzv. *obsazené*. Poté pravděpodobnost, že právě  $m$  infekcí je přeneseno vrcholem stupně  $k$ , zřejmě činní

$$\binom{k}{m} T^m (1-T)^{k-m}. \quad (1.46)$$

Definujme analogicky vytvářející funkci  $\Gamma_0(z, T)$  pro pravděpodobnostní rozdělení počtu obsazených hran pro náhodně vybraný vrchol (jinak řečeno počet infekcí, které způsobil náhodně vybraný jedinec), tj.

$$\Gamma_0(z, T) = \sum_{m=0}^{\infty} \left[ \sum_{k=m}^{\infty} p_k T^m (1-T)^{k-m} \right] z^m. \quad (1.47)$$

Po změně mezí sumace a použití binomické věty dostáváme

$$\begin{aligned} \Gamma_0(z, T) &= \sum_{k=0}^{\infty} p_k \left[ \sum_{m=0}^k (zT)^m (1-T)^{k-m} \right] = \\ &= \sum_{k=0}^{\infty} p_k [zT + (1-T)]^k = G_0(1 + (z-1)T). \end{aligned} \quad (1.48)$$

Stejným způsobem obdržíme vytvářející funkci rozložení přebytkových stupňů vrcholů  $\Gamma_1(z, T)$

$$\Gamma_1(z, T) = G_1(1 + (z - 1)T) \Rightarrow \Gamma'_1(z, T) = TG'_1(1 + (z - 1)T). \quad (1.49)$$

Reprodukční číslo po téhle úpravě můžeme vyjádřit jako

$$\mathcal{R}_0 = \Gamma'_1(1, T) = TG'_1(1). \quad (1.50)$$

Vidíme, že tímto způsobem vyjádřena transmisibilita má přímý vliv na pravděpodobnost vzniku epidemie, respektive úplném zastavení epidemie ( $\mathcal{R}_0 < 1$ ). Je možné také zavést tzv. *kritickou transmisibilitu*  $T_c$ , kde

$$T_c G'_1(1) = 1. \quad (1.51)$$

$T_c$  určuje kritický bod, kdy reprodukční číslo klesne pod 1. Jinými slovy, pokud je transmisibilita snížena pod  $T_c$ , například zvýšenou hygienou či ochrannými pomůckami při kontaktu, je to postačující efekt pro nevznik epidemie.

Je proveditelné určit podmínky, za kterých dojde ke konvergenci výsledků modelu výše, jenž je stochastický, a například modelu SIR? Rozdělme nejdříve poměrově populaci, aby platilo

$$S(t) + I(t) + R(t) = 1. \quad (1.52)$$

Nechť je  $u$  náhodně vybraný vrchol grafu a  $\theta(t)$  značí pravděpodobnost, že náhodný soused nepřenese infekci na  $u$  v čase  $t$ . Nechť  $u$  má  $k$  hran, poté je pravděpodobnost nepřenese infekce v čase  $t$  zřejmě  $[\theta(t)]^k$ . Průměrováním přes všechny vrcholy vyplývá

$$S(t) = \sum_{k=0}^{\infty} P(k) [\theta(t)]^k = G_0(\theta(t)). \quad (1.53)$$

Uvažujme o  $\theta(t)$  jako o součtu třech pravděpodobností:  $\varphi_S$  pravděpodobnost, že náhodný soused je náchylný, respektive  $\varphi_I$ , že je nemocný, ale nepřenese chorobu na  $u$ , respektive  $\varphi_R$  soused je uzdravený. Potom pravděpodobnost nakažení od náhodného souseda musí činit  $(1 - \theta)$ . Nechť hrana má pravděpodobnost, že přenese infekci rovnou  $\beta\varphi_I$ , a dále nechť je rychlost přechodu mezi vyléčené vyjádřena jako  $\alpha\varphi_I$ . Potom tok jedinců mezi těmito kompartmenty můžeme vyjádřit skrz změnu pravděpodobností

$$\varphi'_R = \alpha\varphi_I, \quad \text{resp.} \quad (1 - \theta)' = -\beta\varphi_I, \quad (1.54)$$

neboť jen z  $\varphi_I$  může dojít k uzdravení, respektive k přenosu infekce. Jestliže je velikost populace dostatečně velká, můžeme předpokládat, že každý jedinec udělá

stejný počet kontaktů (značme  $C$ ) za jednotku času. Kombinace tohoto faktu a rovnic (1.53), respektive (1.54) dává

$$\begin{aligned} S(t) &= \theta^C(t); \quad S'(t) = G'_0(\theta(t))\theta'(t) = -G'_0(\theta(t))\beta\varphi_I(t); \\ G'_0(\theta(t)) &= C\theta^{C-1}(t) \quad \Rightarrow \quad S'(t) = -\beta CS(t)\frac{\varphi_I(t)}{\theta(t)} \end{aligned} \quad (1.55)$$

Zavedme konstantu  $\hat{\beta} = \beta C$ , což je reálné pro  $C \rightarrow \infty$ , když  $N \rightarrow \infty$ . Pro velké hodnoty  $C$  je  $\varphi_I$  přibližně  $I$  a  $\theta$  přibližně 1. Jinými slovy platí  $\varphi_I/\theta \approx I$ . Tedy dostáváme

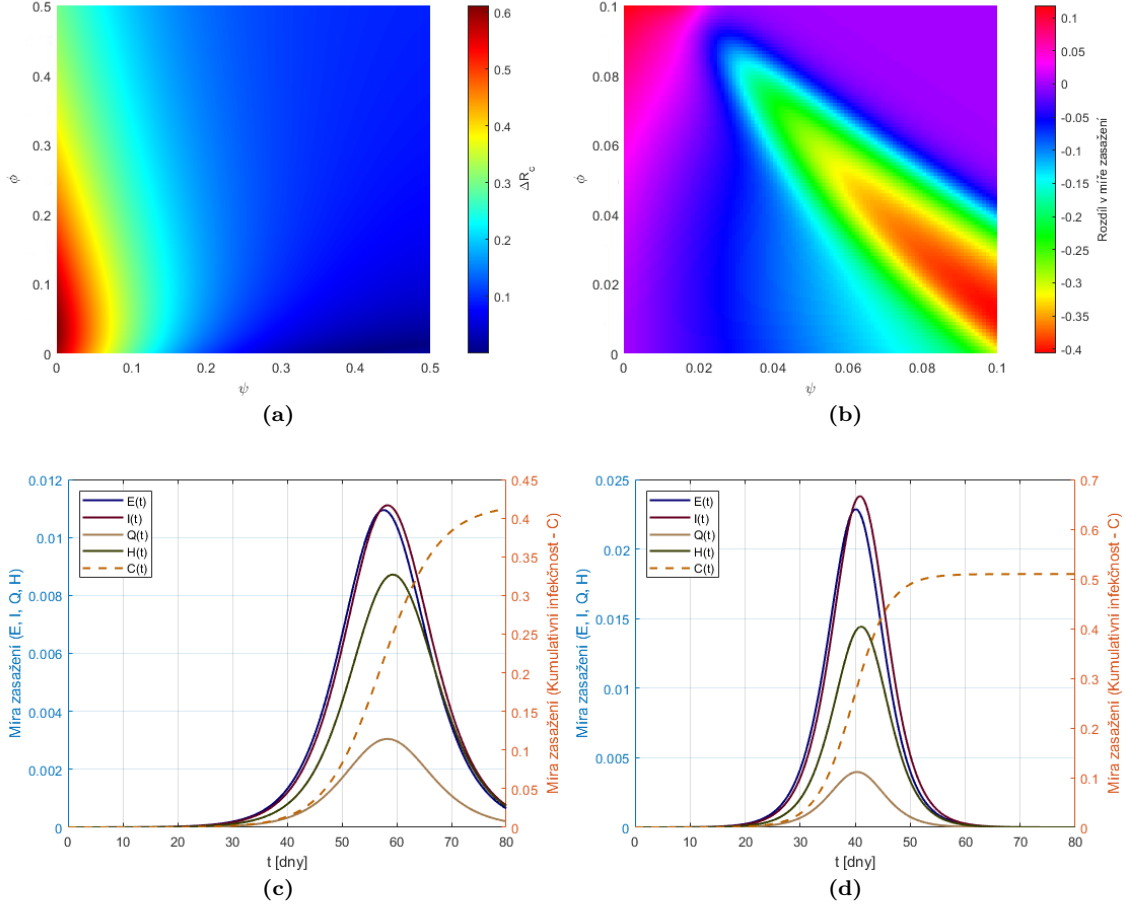
$$S' = -\hat{\beta}SI, \quad (1.56)$$

což je vyjádření první rovnice SIR. Podobně bychom postupovali i u dalších kompartmentů.

Impetus existence této sekce je prvotně odůvodnění některých diskrepancí, které vznikají mezi spojitými deterministickými modely, které předpokládají homogenní mísení obyvatel, a reálnými daty. Grafové modely fungující jako alternativa k čistě stochastickým dynamickým systémům, které také dokáží postihnout dynamiku chování od počátku epidemie. Dalším cílem byla identifikace podmínek za kterých dochází ke konvergenci výsledků kompartmentových a grafových modelů. Jedním z využití těchto *network based* modelů je predikce efektu takzvaných *super-spreader events* (SSEs) na dynamiku epidemie, kdy dochází extrémnímu nárůstu počtu nakažených škálově za velmi krátkou dobu například při sociálních událostech jako koncerty, sportovních utkání, konference nebo proniknutí infikovaného jedince do míst, kde má mnoho jedinců mezi sebou kontakt mnohem vyšší než s okolním světem, jako jsou vězení, domovy důchodců, sanatoria a podobně. Například *Small* a *Tse* [15] analyzují SSE (pomocí grafového modelu) v hotelu Metropol v Hongkongu roku 2003, kde došlo k nákaze 120 osob nemocí SARS od čínského lékaře.

## 1.7 Komparace vybraných modelů

V této sekci uvedme některé numerické simulace, které zvýrazní rozdílnosti v určitých dosud představených kompartmentových modelech, které zatím nebyly zřejmé. Začneme představením změny v chování efektivnosti intervenčních opatření v případě předpokladu exponenciálního a gama rozdělení jedinců ve svých třídách, a to na modelu (1.30), protože podobný bude použit ve třetí kapitole. Budou uvažovány dvě kritéria a to dopad kontrolních opatření (řízeny  $\psi$ , resp.  $\phi$ ) na  $\mathcal{R}_c$  a redukce v konečné velikosti epidemie vyjádřenou jako rozdíl v počtu kumulovaných infikovaných. Pro zjednodušení nechť platí  $\varepsilon_E = 0$ ,  $\kappa_E = \kappa_Q = \kappa$ ,  $\alpha_I = \alpha_H = \alpha$ . Dále je

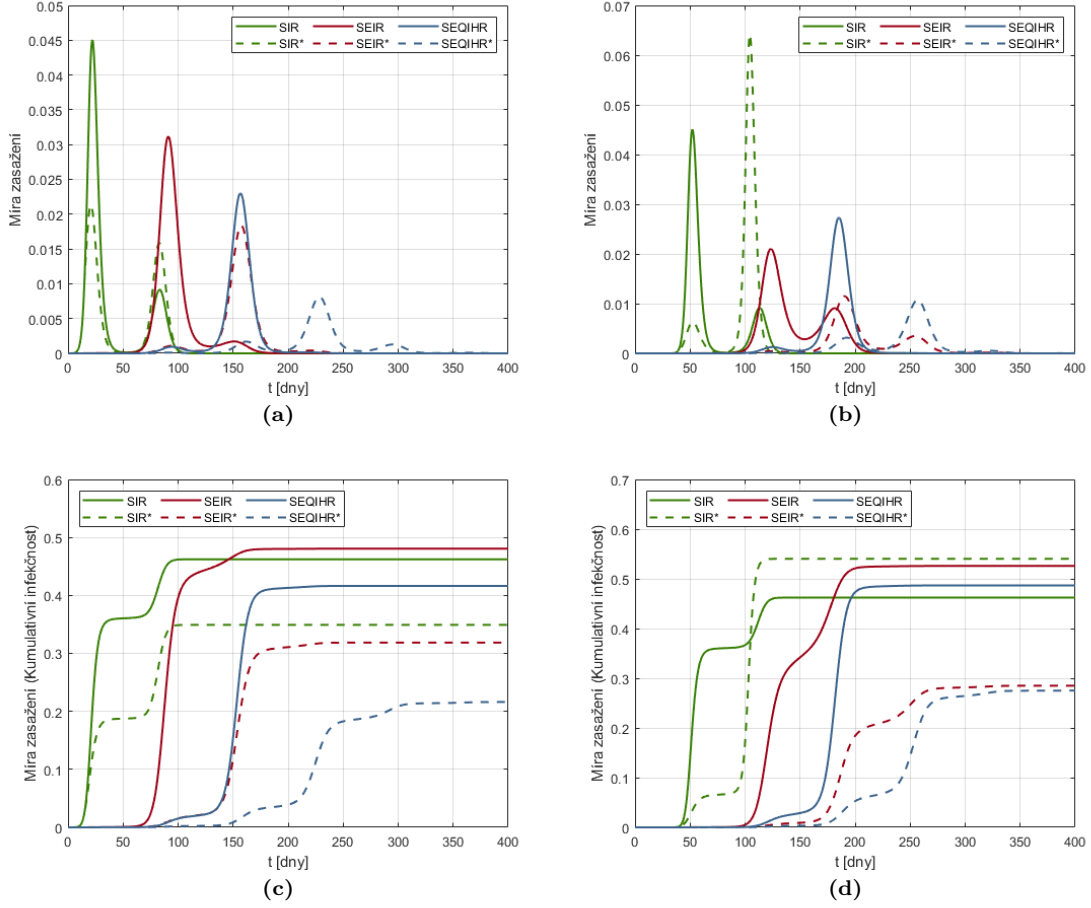


**Obr. 1.5:** (a) simulace rozdílu v  $\mathcal{R}_e$  pro různá  $\phi$  a  $\psi$ . (b) rozdílnost v konečné míře zasažení pro různá  $\phi$  a  $\psi$ . (c) ukázka simulace pro exponenciální distribuci pro  $\phi = 0.07$  a  $\psi = 0.03$ . (d) ukázka simulace pro gama distribuci pro  $\phi = 0.07$ ,  $\psi = 0.03$ ,  $m = 5$ ,  $n = 10$ ,  $N = 10^6$ ,  $I(0) = 1$ ,  $E(0) = Q(0) = H(0) = R(0) = 0$ .

nutné rozšířit (1.32) pro gama rozdělení

$$\mathcal{R}_e = \frac{\beta N}{n\alpha} \left[ 1 - \rho \left( 1 - \left( \frac{m\kappa}{m\kappa + \psi} \right)^m \frac{\alpha}{n\alpha + \phi} \sum_{j=0}^{n-1} \left( \frac{n\alpha}{n\alpha + \phi} \right)^j \right) \right], \quad (1.57)$$

kde  $m$ , resp.  $n$  je počet dělení třídy  $E$  a  $Q$ , resp.  $I$  a  $H$  podobně jako v rovnici (1.33). Necht  $\rho$  představuje  $(1 - \varepsilon_H)$ . Všimněme si  $\partial \mathcal{R}_e / \partial \psi < 0$ , resp.  $\partial \mathcal{R}_e / \partial \phi < 0$ , tedy reprodukční číslo klesá podle očekávání se zpřísňováním opatření. Na Obr. 1.5a vidíme výstup simulace rozdílnosti reprodukčním čísle mezi gama ( $m = 5$ ,  $n = 10$ ) a exponenciálním rozdělením (v tomto pořadí) pro proměnná  $\psi$  a  $\phi$  v intervalu  $[0, 0.5]$ , s parametry  $\kappa = 1/7$ ,  $\beta N = 3/10$ ,  $\rho = 4/5$ ,  $\alpha = 1/10$  (zvoleny, aby přibližně odpovídaly SARS [5]). Jsou patrné signifikantní výkyvy a to zejména v oblastech  $\phi, \psi < 0.1$ , kde lze očekávat hodnoty po zavedení kontrolních opatření [4]. Tedy je vždy nutné zvážit použití daného modelu a vyhodnotit jeho příhodnost, neboť může dojít k velkému zkreslení efektivnosti daných protiepidemických opatření a



**Obr. 1.6:** \* značí aplikace opatření  $\varphi_0$ . (a) simulace SIR, SEIR a SEQIHR po vypuknutí epidemie v  $t_0 = 0$  dní. (c) ten samý případ v podobě kumulace infikovaných. (b) simulace pro  $t_0 = 30$  dní. (d) ten samý případ v podobě kumulace infikovaných. Bylo použito  $m = n = 3$ ,  $N = 10^6$ ,  $I(0) = 1$ ,  $E(0) = Q(0) = H(0) = R(0) = 0$ .

tím vyšší prevalenci choroby. Tento konflikt je možné vysledovat i ve vyobrazení rozdílu v konečné velikosti epidemie (Obr. 1.5b - z důvodu výpočetní náročnosti byl zvolen pouze zmíněný interval a navíc s větším krokem  $\Delta\phi$ ,  $\Delta\psi$ ). Pro hodnoty  $\phi = 0.07$ ,  $\psi = 0.03$  byl model SEQIHR numericky vykreslen na Obr. 1.5c a Obr. 1.5d. Můžeme pozorovat rozdílnost v efektivitě dané strategie, kdy model s gama rozdělením kompartmentů predikuje až o 10% vyšší konečnou zasaženost populace. Kvintesevence diskrepance mezi modely spočívá především v bezpaměťovosti exponenciálního rozdělení [10].

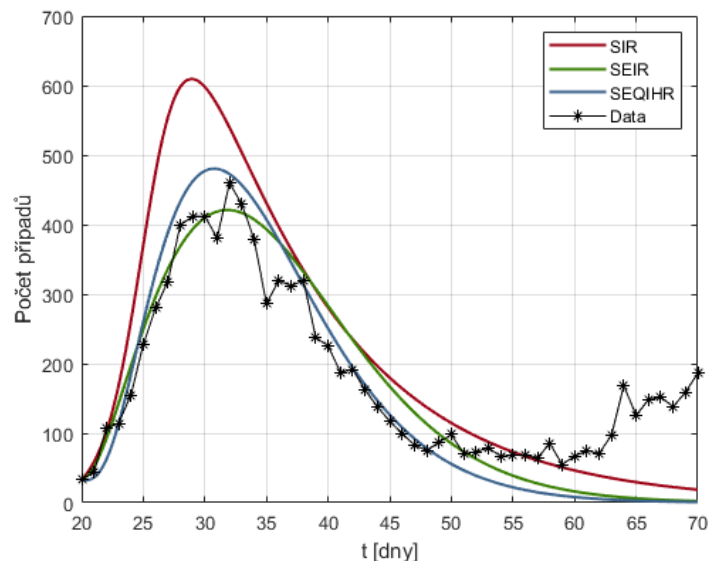
Jako další příklad uvedme simulaci efektu načasování protiepidemických opatření na finální velikost epidemie s předpokladem proměnného parametru  $\beta$ . Nechť platí

$$\beta(t) = \beta_0[1 + \nu \cos(2\pi t/h)], \quad (1.58)$$

kde  $\nu$  představuje velikost fluktuace okolo  $\beta_0$  a  $h$  periodu této fluktuace. Výraz může

být interpretován jako změna chování jedinců během různých časových období (ročních období apod.) či uvolňování a následné zpřísnění kontrolních opatření. Nechť je první jedinec infikován v čase  $t_0$  a dále dejme ať jsou s desetidenním zpožděním aplikovány razantní opatření (zavření škol apod.) či vakcinace, které za velmi krátkou dobu sníží počet náchylných o faktor  $\varphi_0$ . Na Obr. 1.6 jsou vyobrazeny ukázky simulací pro modely (1.4), (1.19), (1.30). Byly zvoleny parametry, které nejvíce odrážejí chování při nákaze virem H1N1 [6] ( $\alpha = 1/3$ ,  $\kappa = 1/2$ ). Dále nechť je SEQIHR upraven jako v případě výše a nechť platí  $\rho = 0.9$ ,  $\psi = \phi = 0.01$ . Pro  $\beta(t)$  určíme  $\beta_0 = 0.5$ ,  $\nu = 0.35$ ,  $h = 90$  dní a ať ad hoc platí  $\varphi_0 = 0.16$ . Z obrázku je možné odpozorovat nuance v chování modelů nejen mezi sebou, avšak i v odpovědi na odlišný okamžik vstupu patogenu do populace (vliv změny ve velikosti (1.58)) a odpovědi modelu na snížení počtu jedinců ve třídě  $S$ . SIR model předpovídá nejrychlejší nástup epidemie a druhou vlnu s nástupem první vlny predikovanou modelem SEIR v obou případech. Dále SIR vykazuje závislost na  $t_0$ , kdy ve druhém případě vidíme přibližně stejnou hodnotu konečné velikosti epidemie, ale při aplikaci  $\varphi_0$  dochází k nástupu velmi výrazné druhé vlny a konečná velikost je až o 20% vyšší než v prvním případě a dokonce vyšší než bez aplikace faktoru  $\varphi_0$ , což se zdá jako paradoxní fenomén. Tento fakt musí být vždy brán v potaz při modelování strategie vakcinace populace a určité hladině prevalence choroby v populaci. SEIR predikuje v obou případech dvě nestejně velké překrývající se vlny epidemie. V konečné velikosti model vykazuje největší zasažení populace bez intervence v podobě  $\varphi_0$ . Model SEQIHR byl zobrazen jako součet  $H(t) + I(t)$ . V obou případech (bez  $\varphi_0$ ) byla predikována relativně velmi malá první vlna epidemie s navazující mnohem závažnější vlnou. V případě intervence  $\varphi_0$  byla vždy celková zasaženost populace nejnižší ze zkoumaných modelů, avšak velmi podobná predikci SEIR u pozdějšího vzniku infekce do populace. Celkově predikce SIR udávají propuknutí epidemie vždy o několik časových jednotě dříve než ostatní dva modely, neboť neobsahuje třídu  $E$  s parametrem  $\kappa$ , který opoždí přechod mezi  $S$  a  $I$ . Při konstantním  $\beta$  by bylo možné předpovídat vždy větší finální velikost epidemie u modelu SEIR než SEQIHR, neboť parametry  $\rho$ ,  $\phi$ ,  $\psi$  mají vliv na velikost kompartmentu  $I$ . To však zřejmě nelze obecně tvrdit u  $\beta(t)$ .

Na závěr uvažujme kazuistiku, první vlnu případů lidské nákazy virem prasečí chřipky (H1N1) v Ciudad de México v roce 2009, kdy se Mexiko stalo ohniskem světové pandemie. Data byla převzata z [16]. Počty zaznamenaných případů na tomto území mezi dvacátým a sedmdesátým dnem epidemie je vykreslen na Obr. 1.7 společně s predikcemi modelů SIR, SEIR a SEQIHR (v podobě, jenž byl představen v této sekci). Průměrná délka infekčnosti byla odhadnuta na  $1/\alpha = 4.1$  dne, dále průměrná doba ve skupině  $E$  na  $1/\kappa = 1.9$  dne,  $\psi = 0.01$ ,  $\phi = 0.04$  a odhad rychlosti prvotního exponenciálního růstu na  $r = 0.22$  [16]. Poté s předpokladem



**Obr. 1.7:** Ciudad de México 2009 - H1N1. 20. až 70. den od zaznamenání prvního případu.

gama distribuce pro kompartmenty ( $m = n = 3$ ) je možné odhadnout rovnicí (1.18) základní  $\mathcal{R}_0 \approx 2.67$ , z kterého můžeme určit  $\beta = 0.65$ . Dále dejme  $N \approx 10^7$ . Konec stochastické fáze epidemie byl observačně určen na dvacátý den od zaznamenání prvního případu, který byl využit jako výchozí stav pro modely ( $I(0) = 33$ ). Vidíme určení kulminace první vlny u všech modelů na interval okolo třicátého dne epidemie. Žádný z modelů nepredikuje nástup druhé vlny od šedesátého dne epidemie, jak je patrné z grafu. To nejspíše souvisí s otevřením některých škol a podniků [16]. Tento fakt by musel být transformován zejména do parametru  $\beta$ , aby bylo možné vystihnout simulací nástup druhé vlny. (Při této simulaci nebylo testováno, zda hodnoty parametrů jsou optimální, tudíž nelze zaručit, že v rámci daného modelu bylo dosaženo kvality non plus ultra pro daná data.)

Je nutné říci, že všechny modely představené v této kapitole jsou pouze zjednodušující strukturou, na které lze vystavět model popisující specifickou infekční chorobu s nutností přizpůsobit charakteristikám dané choroby, jak bylo popsáno v incipitu této práce. I přesto je model vždy idealizací a predikce mohou být zavádějící. Například jsme u modelu (1.30) předpokládali velký pokles infekčnosti jedinců v karanténě a při hospitalizaci z důvodu omezení kontaktu. Avšak při epidemii SARS 2003 způsobily přenosy z hospitalizovaných pacientů na zdravotnický personál extrémní akceleraci epidemie [8].



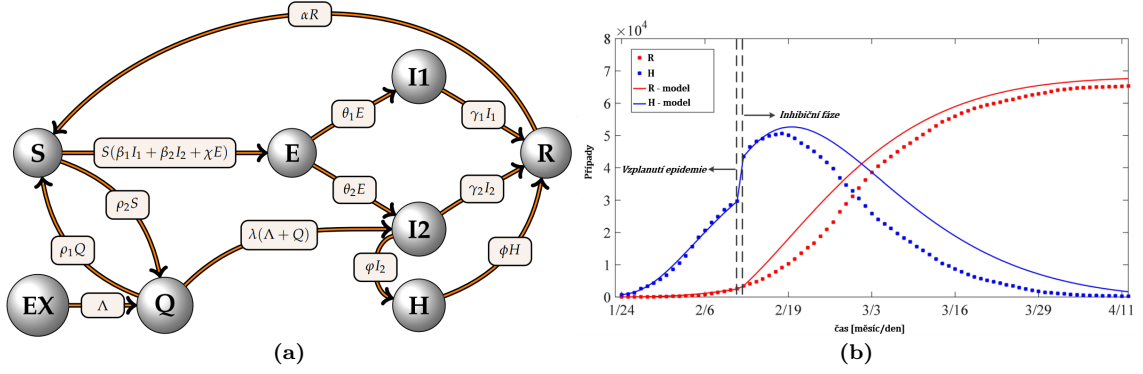
## 2 Rešerše modelování COVID-19

V průběhu roku 2020 vznikly myriády vědeckých prací na téma modelování pandemie COVID-19. Avšak existují velmi odlišné přístupy k tomuto problému, kde deterministické modely tvoří pouze jednu z možností. Dalšími jsou například plně stochastické modely či predikce na základě hlubokého učení, které jsou mnohem přístupnější než dříve díky rostoucí kapacitě výpočetní techniky, či inovativní přístup v podobě analýzy dopravního toku v určité oblasti, jež je založeno na šíření mezi vrcholy grafu. Uvedme některé postupy a závěry vybraných studií relevantních vůči modelům představených v první kapitole.

COVID-19 je respirační onemocnění způsobené virem SARS-CoV-2 (původně nazván 2019-nCoV) se vznikem v pravděpodobně zoonotické události. Pacienta nula lze nejspíše situovat do provincie Chu-pej v Číně do podzimu 2019. První epicentrum pandemie vzniklo ve městě Wu-chan této provincii. Virus je možné detekovat již 1-2 dny v horních cestách dýchacích před nástupem symptomů. Aktivní replikace viru probíhá v průměru 7-12 dní [20].

*De Brouwer et al.* [17] použili jednoduchý SEIR model (1.19) a Bayesovskou regresi pro určení parametrů na data z Belgie, Španělska, Jižní Koreje a pevninské Číny z jarních měsíců 2020, z cílem odhadnout změny  $\mathcal{R}_0$  po oficiálně zavedených kontrolních opatřeních. Autoři odhadli pokles reprodukčního čísla v Číně z původních 4.94 (22.02.2020) na 1.90 (o 69%) a to záhy po pozastavení veškeré veřejné dopravy v postižených regionech. Po uzavření všech zbytných podniků došlo dokonce ke snížení  $\mathcal{R}_0$  podle SEIR modelu na hodnotu 0.055 (znovu se skokovým efektem). To však nelze tvrdit podle autorů o dalších zkoumaných teritoriích, kde vždy docházelo jen k postupnému nástupu efektu (v řádu týdnů). Při fitu modelu na kumulativní počty případů bylo autory dosaženo velmi dobrých výsledků s výjimkou Číny. Autoři dále dokazují bias v testování na příkladu Španělska, kdy z důvodu limitovaného počtu testů a preferovaného použití na symptomatické jedince, došlo k extrémnímu podcenění případů zejména ve věkové skupině 20 – 29 let ( $\approx 7\times$ ). To na základě dat z Jižní Koreji (s demografickou úpravou), kde docházelo intenzivně k plošnému testování obyvatelstva, a údajů o mortalitě. Autoři spekulují o vlivu této věkové skupiny na iniciační fáze epidemie, neboť se jedná o sociálně velmi aktivní skupinu. Dále bylo prokázána neproveditelnost opatření, které by zajišťovala pouze promořování populace, neboť například podle modelu SIR je pro  $\mathcal{R}_0 = 3$  při kulminaci epidemie infekční 30% populace, avšak kapacita lůžek na JIP v EU je pouze přibližně 20 na 100 tis. obyvatel. Takže saturace systému je nevyhnutelná při odhadu 1% nemocných s kritickým průběhem COVID-19 a časová škála takové strategie je několik let.

V článku [18] *He et al.* provádí analýzu svého původního modelu, jenž vychází



**Obr. 2.1:** (a) digram modelu SEQIHR podle [18]. (b) (převzato z [18] - upraveno) simulace modelu pro data z Chu-pej.

ze SEQIHR (1.30), a jeho aplikaci na data z Chu-pej z počátku roku 2020. Diagram modelu je vyobrazen na Obr. 2.1a (jsou použité odlišné značení pro parametry řídící tok mezi kompartmenty). Jsou patrné rozdílnosti vůči (1.30), jako je návrat jedinců zpět do  $S$ .  $I_2$  značí nemocné, jejichž množství kontaktů bylo sníženo na základě intervencí.  $Q$  tentokrát představuje jedince zdravé, avšak v izolaci.  $EX$  značí vnější populaci. Parametry modelu, které nejsou zjistitelné z referenčních biochemických analýz, byly odhaleny pomocí algoritmu optimalizace hejnem částic (*particle swarm optimization*). Autoři dále analyzují podmínky vzniku chaosu v systému na základě algoritmu 0 - 1 testu (*0 - 1 test for chaos*) a předpokladu šumu a periodicity (s řádově malou amplitudou) zejména v koeficientech  $\beta_1$  a  $\beta_2$ . Test dokáže odlišit regulární chování systému (např. periodicitu, kvazi-periodicitu) od chaotické dynamiky. Pracuje přímo s časovou řadou generovanou systémem, tedy není nutné provádět rekonstrukci ve fázovém prostoru [19]. Analýza ukázala, že systém může generovat chaos pro jisté parametry a na určité časové škále. To bylo presentováno na řadě bifurkačních diagramů a příkladů kdy pro určité koeficienty není patrná konvergence pro  $I_1$  či  $I_2$ . Článek však nezkoumal Lyapunův exponent systému, tedy není možné říci, jak přesné musí být určení počátečního stavu epidemie, aby mohla být predikce například po několik týdnů v určitých mezích přesnosti. Na Obr. 2.1b jsou vyobrazeny simulace ze [18]. Odhad velmi dobře kopíruje fázi propuknutí epidemie, ale přesnost velmi klesá ve fázi inhibiční. Autoři tento trend vysvětlují časovou proměnlivostí protiepidemických opatření, která nebyla promítnuta do parametrů, a dále absencí optimalizace více parametrů, neboť část z nich byla převzata z referencí.

Svůj model vycházející také z (1.30) představili *Biswas et al.* [21], který aplikují na první vlnu epidemie v Indii na jaře roku 2020. Model je modifikován přidáním kompartmentu představující asymptomatické jedince, je přidán tok jedinců z vnější populace (parametr  $\pi$ ) a návrat po uzdravení do  $S$ , je brána v úvahu úmrtnost z důvodu nemoci a jiných příčin (parametry  $\delta$ ,  $\mu$ ). Formulace  $\mathcal{R}_0$  byla dovozena po-

mocí generační matice (*next-generation matrix*), konkrétně jako spektrální poloměr matice  $\mathbf{FV}^{-1}$  ( $\mathbf{F}$  vznikla derivací podle jednotlivých členů matice, jež představuje absolutní přírůstek infikovaných za jednotku času, v bodu populace prosté infekce -  $\mathbf{V}$  vznikla podobně, avšak z matice představující toky jedinců mezi kompartmenty [6]). Dále je v článku dokázána existence pozitivní invariantní množiny pro daný model, tj. bez ohledu na podobu atraktoru systému, tak každé řešení, které vstoupí do oblasti hyperkrychle ohraničené čísly 0 a  $\pi/\mu$  pro jednotlivé kompartmenty, nikdy tento region nemůže opustit [14]. Autoři rozdělili data na před a po vládních intervencích. K odhadu parametrů použili metodu nejmenších čtverců. Model mnohem lépe predikoval chování epidemie po zavedení kontrolních opatření. To bylo přičítáno nejistotě v počtu nedetekovaných případů na počátku epidemie.  $\mathcal{R}_0$  v předintervenční fázi epidemie bylo odhadnuto na základě počátečního růstu stejně jako v rovnici (1.18) na hodnotu  $\approx 2.39$ . Ve fázi s protiepidemickými opatřeními autoři určovali takzvané *efektivní reprodukční číslo* ( $\mathcal{R}_{ef}$ ), jež je závislé na aktuálním přírůstku nových případů choroby, respektive přírůstcích od počátků epidemie z jednotlivých časových jednotek, respektive pravděpodobnostním rozdělení setrvání jedinců v jednotlivých kompartmentech. V článku je předpokládána exponenciální distribuce. Pro takto definované reprodukční číslo je možné hodnotit efektivitu kontrolních opatření v čase. Autoři odhadují, že až 45% případů je asymptomatických. Simulace dále ukazují, při nulovém efektu opatření by byla míra zasažení až 37% obyvatel Indie. Pokud však například 77% obyvatel o polovinu sníží riziko přenosu (tj. koeficient  $\beta$ ) například přes redukci denních sociálních kontaktů, vyšší hygienickou bdělostí apod., bude konečná velikost epidemie tvořena jen 0.28% indické populace.

Garba et al. [22] formulují podobně jako Biswas svůj model na základě struktury SEQIHR, s přídatnými kompartmenty pro asymptomatické nakažené, nemocné s mírnými příznaky a nemocné s těžkým průběhem. Je zohledněna míra percepce (ve smyslu změny chování) epidemie mezi obyvatelstvem v podobě skupiny  $P$ , jež je vymezena z  $S$ . Byl umožněn tok jedinců do populace. Tito jedinci mají sníženou pravděpodobnost nákazy. Koeficient  $\beta$ , jež řídí počet nových nemocných, je vyjádřen podle [22] jako

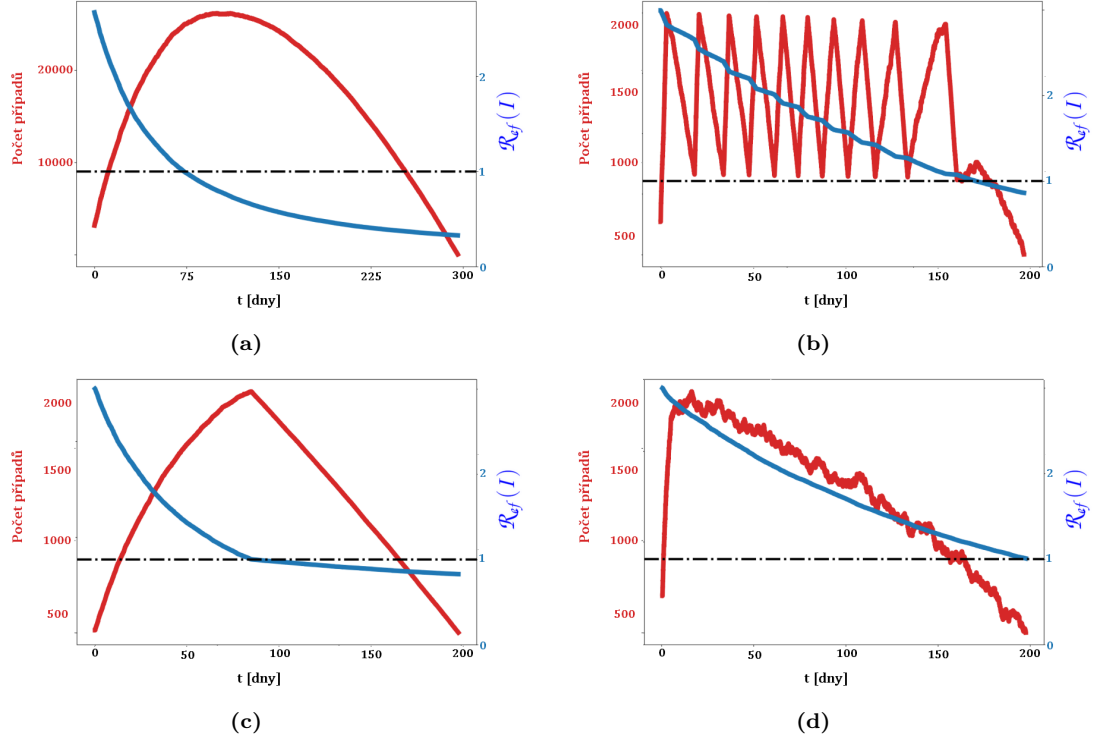
$$\beta(t) = \beta_0(1 - a) \left[ 1 - \frac{P(t)}{N(t)} \right]^\sigma, \quad (2.1)$$

(koeficienty upraveny z důvodu vyhnutí se ambivalenci v textu) kde  $a$  udává efektivnost intervenčních opatření ( $0 \leq a \leq 1$ ),  $\sigma$  určuje odhad síly odpovědi populace na tyto opatření. Reprodukční číslo pro model byl dovozen znovu přes generační matice. Autoři tento model použili na simulaci predikce chování první vlny epidemie v JAR ve druhém a třetím čtvrtletí roku 2020. Byla vyloučena data před omezením

přístupu do země a karanténa pro navrátníci se jedince, aby bylo možné vyloučit příliv nedetekovaných jedinců a tím zamezit chybám v predikcích. V článku je poté provedena simulace efektu lockdownu s trváním 35 dní, který je následován dalším lockdownem o proměnlivé délce někdy v průběhu následujících 70 dní s předpokladem snížení frekvence kontaktů o více než 50%. Bez intervence by činil přírůstek podle predikce až 100 tis. případů denně při kulminaci pandemie na počátku července 2020. S vládními opatřeními (35+21 dní lockdown) byla očekávána kulminace až v druhé polovině července s maximálním přírůstkem mezi 13 – 14 tis. případů denně. Podle [23] došlo k vrcholu pandemie koncem července s maximálním přírůstkem 13944 případů. Avšak oficiální vládní kontrolní opatření měli trvání od začátku května 2020 do konce září 2020 v různých stupních omezení, kdy nejprísnejší byla v platnosti květen až červen. Tedy není zcela možné potvrdit nexus mezi predikcemi a skutečným průběhem. [22] dále udává simulace závislosti efektivnosti detekování nových případů (neboť jejich infekčnost je signifikantně snížena následnou izolací) na konečnou velikost pandemie a ukazuje, pokud dojde k úspěšné detekci již 50% nových nakažených, tak lze očekávat 60% snížení v konečné velikosti zasažení.

Vliv heterogenity v kontrolních strategiích na různé věkové skupiny zkoumá *Yang et al.* [24] s predikcí pro demografická data ze státu São Paulo, Brazílie. Použitý model vychází ze SEQIHR s přidáním skupin asymptomatických jedinců  $A$ , rozdělením kompartmentu  $Q$  na nakažené a zdravé, je umožněn tok jedinců do populace a letalita z důvodu choroby. Model je rozdělen na dvě části pro demografickou skupinu 20 – 50 let, respektive 50 >, kdy tok mezi těmito dvěma skupinami je řízen koeficientem na časové škále odpovídající věkovému rozložení populace. Autoři zkoumali scénáře aplikace izolace pro obě věkové skupiny v různé paritě a následné konsekvence na počty nemocných ve vážném stavu. Například v režimu, kdy polovina obou skupin bude v izolaci (tj. signifikantně sníží svůj počet kontaktů za jednotku času) dojde sice ke zpoždění kulminace epidemie, avšak počty nemocných s těžkým průběhem budou srovnatelné jako v případě scénáře bez izolace. U strategie, kdy pouze starší věková skupina zůstane izolována, dojde ke snížení těžkých případů u této skupiny o přibližně 50% a o  $\approx 30\%$  u všech obyvatel. Autoři připouští, že takový efekt je pro protiepidemická opatření žádoucí, avšak implementace v takového režimu je obtížná.

*Oz et al.* [20] zkoumají vznik kolektivní imunity a takzvaný *after-burn effect* epidemie na datech z různých míst světa. Jedná se o fázi vyhasínání nákazy, kdy se efektivní reprodukční číslo drží pod hodnotou 1, avšak může trvat dlouhou dobu než dojde k úplnému zastavení šíření. Autoři disputují tvrzení, jenž tvoří predikce spojitých deterministických modelů, že ke vzniku kolektivní imunity populace například pro počáteční  $\mathcal{R}_0 = 3$  dochází u parametrů COVID-19, až při nákaze nejméně 40% obyvatel vymezeného území. Takový průběh by byl podle článku pravděpodobný



**Obr. 2.2:** (převzato z [20] - upraveno) after-burn efekt a kolektivní imunita populace pro  $N = 10^6$ ,  $1/\alpha = 10$  dní,  $\mathcal{R}_0 = 3$ .  $\mathcal{R}_{ef}$  je závislé na  $I(n)$  (a) žádná kontrolní opatření. (b) lockdown jen pokud prevalence nákazy překročí práh (0.2% populace v simulaci). (c) jeden lockdown po dosažení  $\mathcal{R}_{ef} \leq 1$ . (d) přísná opatření kopírující závažnost epidemie, avšak vždy jsou pod prahem vstupu do lockdownu.

jen v případě homogenního mísení populace, tj. každý jedinec by měl stejnou pravděpodobnost, že se nakazí v dalším časovém kroku (jak bylo již řečeno v kapitole 1.6). Jinými slovy diskrepance oproti deterministickým modelům jsou v tomto případě signifikantní a je nutné je vysvětlit. Autoři článku používají k modelování graf kontaktů jedinců. Předpokládají heterogenitu v počtu kontaktů v podobě různých distribucí pravděpodobností nákazy. Tedy existují nemocní jedinci, kteří mají mnohem větší pravděpodobnost dalšího šíření, ale také zdraví jedinci s mnohem větší šancí stát se nakaženými než ostatní. Právě korelace mezi těmito skupinami a existence superpřenašečů má podle autorů drastický efekt na odhad procenta populace pro získání kolektivní imunity a také nadsazení  $\mathcal{R}_{ef}$  i o desítky procent než ve skutečnosti. Nejdříve jsou dovozeny výsledky pro obecnou podobu pravděpodobnostních rozdělení nakažlivosti pro jednotlivé uzly grafu. Konkrétní simulace jsou s použitím gama rozdělení a Paretova rozdělení. Za předpokladu gama distribuce je dospěno ke vztahu vyjadřující míru zasažení populace, kdy je dosaženo kolektivní imunity,

$$\varepsilon = 1 - \mathcal{R}_0^{-\frac{\alpha}{\alpha+2}} \quad (2.2)$$

(podle [20]). Tedy například pro  $\mathcal{R}_0 = 3$  a  $1/\alpha = 6$  dní dostáváme  $\varepsilon \approx 9\%$ , což je

pokles o 30% oproti deterministickému modelu. Dále je kvantitativně vyhodnocen after-burn efekt, kdy i po dosažení kolektivní imunity populace může počet nových případů klesat jen pozvolna. Na Obr. 2.2 jsou grafické výstupy těchto simulací, kde kolektivní imunita je dosažena při poklesu pod  $\mathcal{R}_{ef} = 1$ . Výsledky [20] implikují nutnost ponechání lockdownu v platnosti, respektive zavést lockdown v případě dosažení kolektivní imunity pro zamezení after-burn efektu. Jak je patrné z Obr. 2.2a dochází k extenzivnímu šíření nemoci dlouho po promoření populace. V případě reakčního přechodu do lockdownu po překročení prahu zasažení a zpět do mírných kontrolních opatření je možné vidět na Obr. 2.2b. Počet nových případů strmě klesá po dosažení kolektivní imunity. Populace na Obr. 2.2c zavádí lockdown ihned po dosažení  $\mathcal{R}_{ef} = 1$ , vidíme mnohem rychlejší pokles než v prvním případě. Čtvrtá populace zavádí opatření, které nejsou tak přísná jak lockdown, avšak drží stále hodnoty nových případů pod vymezeným prahem. V tomto případě dosažení kolektivní imunity nastává téměř ve stejném okamžiku, kdy se populace stává prostou infekce.

### 3 Design modelu a simulace

Jak bylo nastíněno v předchozích dvou kapitolách, existuje mnoho přístupů pro vystižení trendu probíhající pandemie. Konkretizujme snahy na oblast Itálie první poloviny roku 2020 pro celou kapitolu. Některé studie jako [29] adoptují jednoduchý SEIR model, které disponují výhodou malého počtu parametrů, avšak za cenu velkého poklesu schopnosti sledovat dynamiku některých skupin infikovaných. Na druhou stranu například autoři [28] integrují do struktury svého modelu téměř state of the art biologické charakteristiky COVID-19 v podobě mnoha infekčních skupin a definic toků mezi nimi. Avšak je zřejmé, že není vždy možné či později proveditelná validace v takovém případě predikce s konkrétními daty. Toto úskalí bude pivotním prvkem ve formulaci původního modelu.

#### 3.1 Model

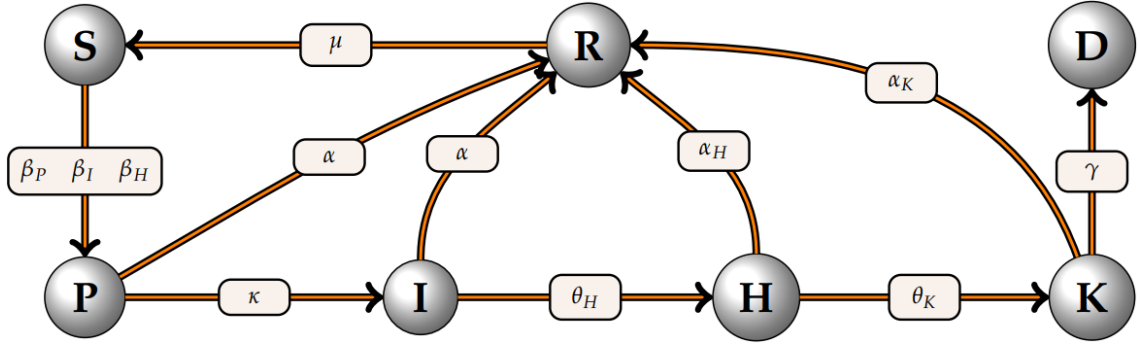
Uvažujme model, který bude odrážet nejlépe kategorie nemocných bez dodatečného dovozování, tj. takové, které byly monitorovány u COVID-19 v průběhu první vlny pandemie. Jinými slovy kalibrace parametrů by měla probíhat na základě dat pro co nejvíce kompartmentů. Nechť je tedy model definován jako

$$\begin{aligned} S' &= -\frac{S}{N}(\beta_P P + \beta_I I + \beta_H H) + \mu R \\ P_1' &= \frac{S}{N}(\beta_P P + \beta_I I + \beta_H H) - m(\kappa + \alpha)P_1 \\ P_j' &= m(\kappa + \alpha)P_{j-1} - m(\kappa + \alpha)P_j, \quad j = 2, \dots, m \\ I_1' &= m\kappa P_m - (\theta_H + m\alpha)I_1 \\ I_j' &= m\alpha I_{j-1} - (\theta_H + m\alpha)I_j, \quad j = 2, \dots, m \\ H' &= \theta_H I - (\theta_K + \alpha_H)H \\ K' &= \theta_K H - (\alpha_K + \gamma)K \\ D' &= \gamma K \\ R' &= m\alpha(I_m + P_m) + \alpha_H H + \alpha_K K - \mu R, \end{aligned} \tag{3.1}$$

kde

$$P = \sum_{j=1}^m P_j, \quad I = \sum_{j=1}^m I_j. \tag{3.2}$$

Tedy nechť skupina  $S$  znovu označuje jedince náchylné k infekci.  $P$  je počet jedinců, jenž jsou symptomatictí či asymptomatictí. Je zřejmé, že každý nemocný jedinec musel projít touto skupinou, neboť detekce nastává až po nakažení.  $I$  značí nakažené, detekované, a/symptomatické jedince a nechť mají tito jedinci sníženou (nebo pozměněnou) schopnost šíření infekce například z důvodu karantény.  $H$  nechť jsou



Obr. 3.1: Diagram modelu (3.1) - (pro přehlednost vyobrazeny pouze parametry toku).

hospitalizovaní jedinci avšak v nekritickém stavu,  $K$  jedinci v kritickém stavu (na JIP),  $D$  počet zemřelých z důvodu choroby a  $R$  uzdravení. Nechť je celková počáteční populace zachována, tj.  $N = S + P + I + H + K + D + R$ ,  $N' = 0$  pro všechna  $t \geq 0$ .  $\beta_P$ ,  $\beta_I$ ,  $\beta_H$  značí množství kontaktů za jednotku času mezi členy  $S$  a  $P$ , respektive  $I$ , respektive  $H$ . Aby byly tyto koeficienty nezávislé na velikosti populace, jsou vyjádřeny v systému dělení  $N$  na rozdíl od první kapitoly. Pro zjednodušení nepředpokládáme přenos (či je řádově zanedbatelný) choroby z pacientů na JIP a zpětný tok pacientů  $K \rightarrow H$ , respektive  $H \rightarrow I$ .  $\kappa$  řídí úspěšnost detekce nemocných.  $\theta_H$ , respektive  $\theta_K$  určují množství jedinců, u kterých dojde ke zhoršení průběhu onemocnění a musí dojít k hospitalizaci, respektive je nutné je umístit na jednotku intenzivní péče.  $\gamma$  vymezuje jedince, kteří chorobě podlehnou, avšak předpokládejme takový scénář jen pro skupinu  $K$  (což je přirozené před saturací systému zdravotní péče daného regionu).  $\alpha$ ,  $\alpha_H$ ,  $\alpha_K$  řídí rychlost uzdravování (přechod do neinfekční fáze). Byly předpokládány odlišné hodnoty pro kompartmenty  $I$ ,  $H$ ,  $K$ , avšak stejné pro  $I$  a  $P$ , protože nebyla dokázána změna doby infekčnosti u COVID-19 u symptomatických (mírný průběh) a asymptomatických jedinců [17], ale existuje u pacientů s těžším průběhem onemocnění. Dále je umožněn tok  $R \rightarrow S$ , řízen parametrem  $\mu$ , neboť byly potvrzeny případy druhotného prodělání nákazy [17]. Diagram modelu je na Obr. 3.1. V meritu sekce 1.7 bylo zvoleno gama rozdělení setrvání ve skupinách  $P$  a  $I$ , které je řízeno tvarovým parametrem  $m$ , a tedy bylo nutné zavést vztah (3.2) pro první, respektive druhou rovnici (3.1) a vznik nových nález. Role některých parametrů může být viděna i v podobě pravděpodobností. Například zřejmě  $\frac{\alpha}{\alpha + \theta_H}$  udává pravděpodobnost zotavení detekovaného jedince bez toho aniž by byl hospitalizován, podobně  $\frac{\alpha_H}{\alpha_H + \theta_K}$  představuje šanci hospitalizovaného na vyléčení bez přesunu na JIP,  $\frac{\gamma}{\gamma + \alpha_K}$  pravděpodobnost úmrtí jedince na JIP či  $\frac{\kappa}{\alpha + \kappa}$  pravděpodobnost detekce infikovaného jedince.

Již v první kapitole byla popsána krucální role reprodukčního čísla při popisování dynamiky epidemie daným modelem. Bylo také patrné, že dovození  $\mathcal{R}_0$  pro tři a více rozměrné modely v podobě hledání parametrů stability ve vlastních číslech jakobiánu



okolo kritického bodu systému je velmi pracné. Jak bylo naznačeno ve druhé kapitole, tento postup lze alternativně provést přes generační matice a již klasický výsledek *Plemmonse* a *Bermana* [26]. Nechť nezáporný vektor  $\mathcal{F}$  představuje tok nových infekcí v populaci, respektive nezáporný vektor  $\mathcal{V}$  (tok násobíme  $-1$ ) v každém svém prvku tok z jednotlivých kompartmentů. Poté nechť je každý prvek  $\mathcal{F}$ , respektive  $\mathcal{V}$  linearizován okolo bodu populace prosté infekce (značme  $y_0$ ), jenž dá vzniknout  $n \times n$  maticím  $\mathbf{F}$ , respektive  $\mathbf{V}$

$$\mathbf{F}_{i,j} = \frac{\partial \mathcal{F}_i}{\partial x_j}(y_0), \quad \text{resp.} \quad \mathbf{V}_{i,j} = \frac{\partial \mathcal{V}_i}{\partial x_j}(y_0), \quad (3.3)$$

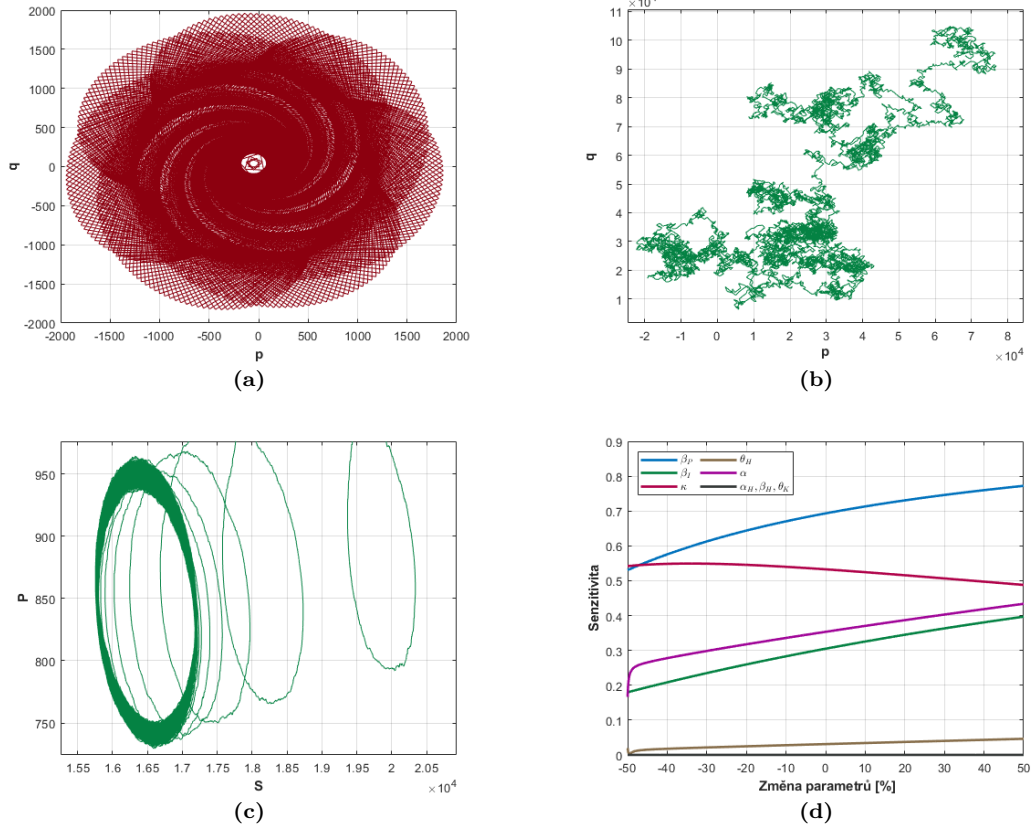
kde  $i$  značí  $i$ -tý prvek vektoru  $\mathcal{F}$ ,  $\mathcal{V}$ , respektive  $j$  představuje  $j$ -tý kompartment, respektive  $n$  počet kompartmentů. V našem případě ovšem  $n - 2$  neboť první rovnici (3.1) vynecháváme, protože může být vyjádřena ze zbylých rovnic z důvodu zachování  $N$  a dále tok do  $D$  je pouze jednosměrný a obdobně vyjádřen z rovnice pro  $K$ . Tedy konkrétně pro systém (3.1) dostáváme

$$\mathcal{F} = \begin{bmatrix} \frac{S}{N}(\beta_P P + \beta_I I + \beta_H H) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \Rightarrow \mathbf{F} = \begin{bmatrix} \overbrace{\beta_P \dots \beta_P}^m & \overbrace{\beta_I \dots \beta_I}^m & \beta_H & 0 & 0 \\ 0 \dots 0 & 0 \dots 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 \dots 0 & 0 \dots 0 & 0 & 0 & 0 \end{bmatrix}, \quad (3.4)$$

kde  $\mathcal{F}$  je  $(2m + 3)$ -prvkový vektor a  $\mathbf{F}$  má dimenzi  $(2m + 3) \times (2m + 3)$ . Podobně platí

$$\mathcal{V} = \begin{bmatrix} m(\kappa + \alpha)P_1 \\ \vdots \\ m(\kappa + \alpha)(P_m - P_{m-1}) \\ (\theta_H + m\alpha)I_1 - m\kappa P_m \\ \vdots \\ (\theta_H + m\alpha)I_m - m\alpha I_{m-1} \\ (\theta_K + \alpha_H)H - \theta_H I \\ (\alpha_k + \gamma)K - \theta_K H \\ \mu R - m\alpha(P_m + I_m) - \alpha_H H + \alpha_K K \end{bmatrix} \Rightarrow \mathbf{V}_1^{-1} = \begin{bmatrix} \frac{1}{m(\kappa + \alpha)} \\ \vdots \\ \frac{\kappa}{(\kappa + \alpha)(m\alpha + \theta_H)} \\ \vdots \\ \frac{\kappa(m\alpha)^{m-1}}{(\kappa + \alpha)(m\alpha + \theta_H)^m} \\ \frac{\kappa\theta_H(2m\alpha + \theta_H)}{(\kappa + \alpha)(\alpha_H + \theta_K)(m\alpha + \theta_H)^2} \end{bmatrix}, \quad (3.5)$$

kde  $\mathcal{V}$  je  $(2m + 3)$ -prvkový vektor a  $\mathbf{V}_1^{-1}$  již značí první sloupcový vektor inverzní matice  $\mathbf{V}$ , neboť hledáme spektrální poloměr matice  $\mathbf{FV}^{-1}$ .  $\mathbf{V}^{-1}$  byla vyjádřena z  $\mathbf{V}$  pomocí Gauss-Jordanovy eliminace převedením na trojúhelníkový tvar, sledování znaménka permutací u eliminace, využití invariance determinantu matice při odečtení násobku jednotlivých vektorů mezi sebou a nakonec vtažení pro mezi inverzní maticí a maticí adjungovanou a determinantem matice (důkaz například v [31]). Z



**Obr. 3.2:** (a) výstup 0-1 testu v  $pq$  rovině s prvním nastavením (viz text). (b) výstup 0-1 testu v  $pq$  rovině s druhým nastavením (viz text). (c) řez podivným atraktorem v  $SP$  rovině pro druhé nastavení 0-1 testu. (d) citlivost (absolutní hodnoty)  $\mathcal{R}_0$  na jednotlivých parametrech a při vychýlení od  $\mathcal{R}_0 = 2$ .

důvody přehlednosti byl vyjádřen pouze první sloupcový vektor  $\mathbf{V}^{-1}$  ve (3.5) a navíc bez dvou posledních prvků neboť při vynásobení  $\mathbf{F}$  s maticí  $\mathbf{V}^{-1}$  budou nulové. Je jasná důležitost pouze tohoto vektoru z důvodu vytvoření jediného nenulového prvku na diagonále  $\mathbf{F}\mathbf{V}^{-1}$ , což je hledané vlastní číslo a také spektrální poloměr matice, tedy dostáváme reprodukční číslo pro systém (3.1) v podobě

$$\mathcal{R}_0 = \frac{1}{\kappa + \alpha} \left[ \beta_P + \frac{\kappa\beta_I}{(m\alpha + \theta_H)} \sum_{j=0}^{m-1} \left( \frac{m\alpha}{m\alpha + \theta_H} \right)^j + \frac{\kappa\theta_H\beta_H(2m\alpha + \theta_H)}{(\alpha_H + \theta_K)(m\alpha + \theta_H)^2} \right]. \quad (3.6)$$

Všimněme si nezávislosti vůči parametrům  $\mu$ ,  $\alpha_K$ ,  $\gamma$ .

Často je výhodné kvantifikovat vliv parametrů modelu na reprodukční číslo. Při použití metody podle [25] zavedme senzitivitu parametrů jako

$$\frac{\partial \mathcal{R}_0}{\partial q_i} \frac{q_i}{\mathcal{R}_0}, \quad (3.7)$$

kde  $q_i$  je nějaký parametr systému, jehož vliv zkoumáme. Na obr. 3.2d vidíme numerické vyhodnocení, kdy jako základní hodnota byla uvažováno  $\mathcal{R}_0 = 2$  pro

$[\beta_P, \beta_I, \beta_H, \kappa, \alpha, \alpha_H, \theta_H, \theta_K] = [0.264, 0.026, 0.006, 0.16, 0.030, 0.069, 0.009, 0.024]$  a  $m = 5$  a konkrétní hodnoty odpovídají při změně parametrů 0%. Dále je možné na témže vyobrazení vidět měnící se senzitivitu při proměnné hodnotě v rozmezí  $\pm 50\%$  od základní hodnoty. Významnost senzitivity parametrů tkví ve vlivu, který mají na změnu  $\mathcal{R}_0$ , tj. čím vyšší senzitivita (absolutní hodnota), tím vyšší  $\Delta \mathcal{R}_0$  při dané změně parametrů. Takto vytypované parametry jsou klíčové pro vývoj epidemie podle zvoleného modelu. Z obr. 3.2d je patrné, že jsou to především  $\beta_P$ ,  $\beta_I$ ,  $\kappa$  a  $\alpha$ .

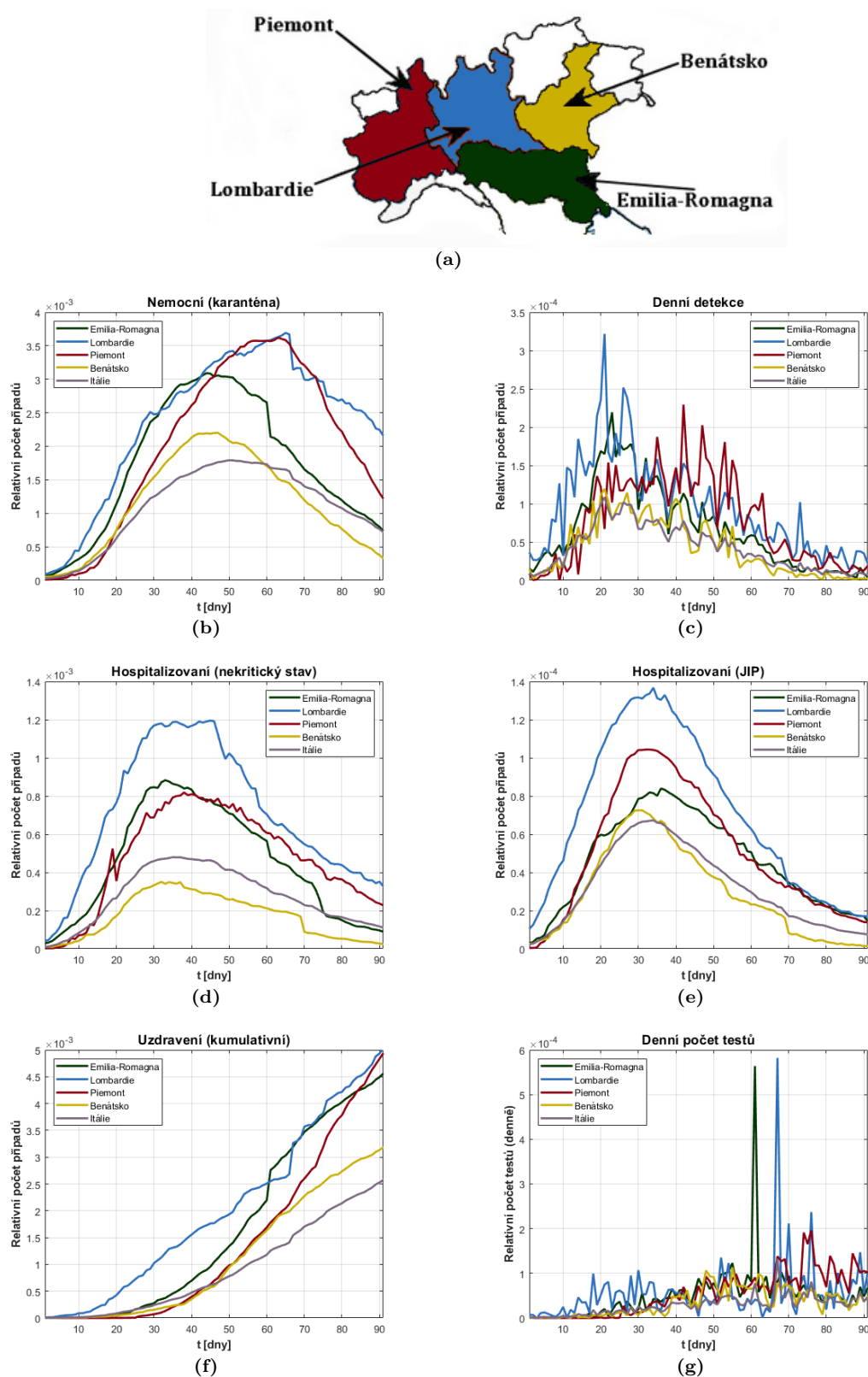
Zkoumání dynamiky a topologie systému (3.1) je z důvodu mnoha rozměrů fázového prostoru ( $> 3$ ) a možnosti neautonomity systému (můžeme připustit závislost parametrů na čase) obtížné studovat analyticky. Omezme snažení pouze na konstatování, zda je možné za určitých podmínek očekávat chaotické chování systému či je dynamika vždy regulární. K tomuto vyhodnocení bude použit 0-1 test podle [19]. Mějme jednorozměrnou časovou řadu generovanou systémem  $\phi(n)$  pro  $n = 1, 2, \dots$ . Využijme například  $P(t)$  s frekvencí 10 vzorků za jednotku času. Dále definujeme podle [19]

$$p(n) = \sum_{j=1}^n \phi(j) \cos(jc), \quad q(n) = \sum_{j=1}^n \phi(j) \sin(jc), \quad (3.8)$$

kde  $c \in (\pi/5, 4\pi/5)$ . Pokud má systém regulární dynamiku, tak by křivka  $(p(n), q(n))$  měla být omezena, v případě chaotického chování by tato křivka v  $pq$  rovině měla mít difuzní tvar (podobný Brownovu pohybu) a tedy pro  $n \rightarrow \infty$  neomezená. Pro simulaci byly užity hodnoty  $[N, P, I, H, K, D, R] = [10^5, 1000, 100, 25, 0, 0]$  s parametry  $[\beta_H, \alpha, \alpha_H, \alpha_K, \theta_H, \theta_K, \gamma, m, \mu] = [0.006, 0.30, 0.069, 0.7, 0.009, 0.024, 0.001, 5, 0.33]$ . Bylo zkoumáno několik případů. V první simulaci bylo chování  $\beta_P$  řízeno podle rovnice (1.58) s  $\beta_{P0} = 5$ ,  $h = 10$ ,  $\nu = 0.8$  s konstantními  $\beta_I = 0.5$ ,  $\kappa = 0.4$  a  $c = 3\pi/10$ . Křivka v rovině  $pq$  však je zřejmě symetrická a ohraničená, jak je patrné z výstupu na obr. 3.2a. Ve druhé simulaci byl testován případ, kdy  $\beta_P$  (zůstává nezměněna) a  $\kappa$ , respektive  $\beta_I$  jsou také proměnné podle (1.58) s  $\beta_{I0} = 0.4$ ,  $h_{\beta_I} = 7$ ,  $\nu_{\beta_I} = 0.3$ , respektive  $\kappa = 5$ ,  $h_{\kappa} = 5$ ,  $\nu_{\kappa} = 0.9$ . Na obr. 3.2b je patrná trajektorie podobná Brownovu pohybu. Na obr. 3.2c je vyobrazen řez podivným atraktorem v fázové  $SP$  rovině pro zmíněné hodnoty parametrů a počáteční podmínky. Je zde viditelná i transientní část řešení konvergující k atraktoru. Tedy lze tvrdit, že dynamika (3.1) může vykazovat chaotické chování a to v určité míře předjímá snížení důvěry v prediktivní schopnost modelu, avšak je nutné podotknout, že hodnoty použité ve druhé simulaci měli extrémně oscilační charakter, který nelze očekávat v reálné situaci.

## 3.2 Data

Pro retrospektivní analýzu a testování (3.1) byla vybrána první vlna pandemie COVID-19 v severní Itálii s daty v rozmezí od 1. března 2020 do 31. května 2020.



**Obr. 3.3:** Grafické vyobrazení některých monitorovaných hodnot z [30]. Den 1 představuje 1. březen 2020. (Jsou důležité pouze diskrétní hodnoty v jednotlivých dnech, zobrazení je spojitě pouze pro přehlednost, spojnice mezi hodnotami nemá žádného smyslu). Hodnoty jsou normalizovány počtem obyvatel regionů, aby byly srovnatelné. (a) regiony zájmu. (b) detekování nemocní (mírný průběh). (c) denní detekce. (d) hospitalizovaní - (nekritický stav). (e) hospitalizovaní na JIP. (f) uzdravení v kumulativní formě. (g) denní počet testů.

Tento interval byl zvolen tak, aby obsahoval pokud možno fázi exponenciálního růstu nakažených a částečně i inhibiční fázi u všech sledovaných regionů. Surová data byla převzata z [30], kam byla postupně nahrávána italským Úřadem pro Civilní Ochranu (*Dipartimento della Protezione Civile*). Pro kalibraci modelu budou využity data pro skupiny  $I$ ,  $H$ ,  $K$ ,  $D$ , protože pouze u těchto dat je možné tvrdit či porovnat, že se jedná o bijektivní vztah mezi jedinci v kompartmentu a zaznamenanými reálnými počty. Monitoring uzdravených jedinců nebude použit pro kalibraci parametrů, neboť předpokládáme velké množství nedetekovaných, tedy by tyto data byla velmi zkreslující. Na obr. 3.3 je možné spatřit graficky některé sledované kategorie jedinců. Obr. 3.3g byl zahrnut, aby mohlo být alespoň observačně řečeno, že nedochází k významné korelaci mezi množstvím testů provedených denně a množstvím nově detekovaných. (V originálních datech byla tato kategorie v kumulačních číslech, proto byl vektor těchto hodnot násoben diferenční maticí a tím byly získány počty na obr. 3.3g). Dva výrazné peaky na zmíněném grafu jsou dva denní pokusy o odhad prevalence a komunitní imunity na COVID-19 (tj. testování i zdravých či asymptomatických jedinců) v populaci vybraných měst v regionech Emilia-Romagna a Lombardie [27].

V další sekci bude důležité sladit kalibraci parametrů modelu s klíčovými regulačními událostmi během vymezeného období epidemie (1.3. - 31.5. 2020), které mohou významně ovlivnit dynamiku šíření. Na základě zavedených kontrolních opatření italské vlády uvedená v [27] byla data rozdělena chronologicky do čtyř fází:

<i>Fáze 1</i>	-	1.3. - 16.3.	-	počáteční fáze bez jednotných kontrolních opatření
				celostátní pozastavení všech neesenciálních služeb a
<i>Fáze 2</i>	-	17.3. - 27.3.	-	komerčních aktivit, zákaz cestování mezi regiony
				severní Itálie
				zprísnění opatření, celostátní lockdown, zákaz i
<i>Fáze 3</i>	-	28.3. - 30.4.	-	sportovních aktivit venku, zrušení mnoha
				vyjímek z minulých nařízení
<i>Fáze 4</i>	-	1.5. - 31.5.	-	graduální uvolňování restrikcí během května
				v závislosti na situaci v jednotlivých regionech

Je nutné podotknout, že zmíněné intervaly byly zvoleny značně heuristicky a tedy při podrobnějším studování lokálních nařízení by mohlo být dospěno k odlišným fázím. Avšak při jemnějším dělení intervalu může dojít k přílišnému snížení poměru mezi počtem reálných hodnot a počtem parametrů modelu a tím ke snížení schopnosti modelu vystihnout dlouhodobější trend.

Pokud byla data v této kapitole normalizována, tak bylo učiněno přes počty obyvatel jednotlivých regionů, respektive celé Itálie uvedené v [30].

### 3.3 Odhad parametrů

Odhad parametrů modelu přes existující data je nezbytné pro reprodukování chování epidemie do minulosti a pro produkci krátkodobých předpovědí. Formálněji nechť  $\vec{Y} = \vec{Y}(t, \vec{p})$  představuje řešení (3.1) pro nějaké  $\vec{Y}_0$ , aby byla zdůrazněna závislost řešení na daném vektoru parametrů. Dále nechť je  $\vec{p} \in \mathcal{P}$ , kde  $\mathcal{P}$  představuje vektorový prostor s  $\dim(\mathcal{P}) = n_p$ , jenž vyjadřuje počet kalibrovaných parametrů násobený počtem fází z minulé sekce. Avšak u specifických parametrů nebudeme očekávat významnou změnu na každé fázi. Jmenovitě určíme  $\alpha$ ,  $\alpha_H$ ,  $\alpha_K$ ,  $\mu$ , u nichž nelze očekávat změnu jako reakci na protiepidemická opatření, tedy budou optimalizovány pouze jednou pro celé sledované období. Nechť dále pro zjednodušení platí u všech fází  $\beta_P \approx c_P \beta_I$ , kde  $c_P$  je konstanta úměrnosti, tj. bude odhadováno právě toto číslo nikoliv  $\beta_I$ . Potom dostáváme  $\dim(\mathcal{P}) = 4 \cdot 6 + 4 = 28$  (bylo vynecháno  $\mu$ , viz další sekce). Je jasné, že předpokládáme konstantní parametry na jednotlivých fázích, tudíž není postulována žádná časová závislost parametrů.

Poté problém můžeme formulovat následovně. Cílem je nalézt řešení tohoto minimalizačního problému

$$\hat{p} = \arg \min_{\vec{p} \in \mathcal{P}} \{ \mathcal{J}(\vec{p}) \}, \quad (3.9)$$

kde  $\hat{p}$  je hledaný optimální vektor parametrů, respektive  $\mathcal{J}(\vec{p})$  je chybová funkce vyjadřující míru odchylky modelu od reálných dat přes různé kompartmenty (konkrétně jsou to již zmíněné  $I$ ,  $H$ ,  $K$ ,  $D$ ). Symbolicky definována jako

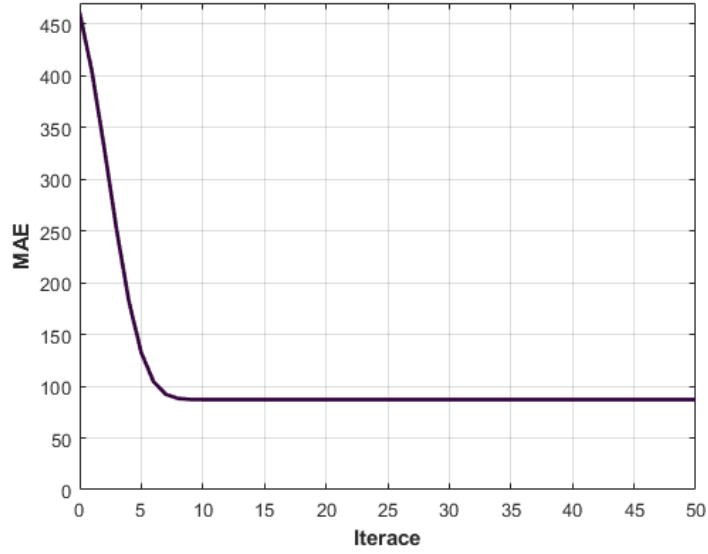
$$\mathcal{J}(\vec{p}) = \sum_{k=\{I,H,K,D\}} \sum_{j=0}^{d-1} \frac{\|\vec{Y}_k(t_j, \vec{p}) - \hat{Y}_k(t_j)\|^2}{\hat{Y}_k(t_j)}. \quad (3.10)$$

$\|\cdot\|$  je klasická Euklidovská norma vektoru.  $\vec{Y}_k$  značí komponenty  $\vec{Y}$  odpovídající zmíněným kompartmentům.  $\hat{Y}_k$  představuje reálné data daných monitorovaných kategorií. Dále tento člen ve jmenovateli zlomku (3.10) působí jako dynamický váhový koeficient (zvoleno podle [11]).  $d$  je počet měření (v našem případě zřejmě  $d = 92$ ).

Prohledávání prostoru  $\mathcal{P}$  za účelem nalezení  $\hat{p}$  bude implementován *Broydenův–Fletcherův–Goldfarbův–Shannonův algoritmus* (BFGS) (celý postup podle [32]). U klasické metody gradientového sestupu je nutné určovat koeficient učení a dochází pouze k využití lineární informace o chování funkce okolo daného bodu. U Newtonové metody je využíváno i kvadratického členu Taylorova rozvoje dané funkce a odpadá nutnost koeficientu učení. Tedy lze iterační krok zapsat

$$\vec{x}_{k+1} = \vec{x}_k - \mathbf{H}^{-1}(\vec{x}_k) \nabla f(\vec{x}_k). \quad (3.11)$$

$\vec{x}_{k+1}$  je nový odhad optimálního řešení,  $f$  optimalizovaná funkce a  $\mathbf{H}^{-1}$  inverzní Hessova matice v bodě  $\vec{x}_k$ . Avšak je nutné v každém kroku tuto matici aktualizovat,



**Obr. 3.4:** Ukázka konvergence MAE (*mean absolute error*) pro data z Lombardie v závislosti na iteracích u implementovaného algoritmu BFGS.

kdy časová složitost této operace je řádu  $\mathcal{O}(n^2)$ , a další je výpočet inverzní matice se složitostí  $\mathcal{O}(n^3)$ . Z tohoto důvodu byly vyvinuty takzvané kvazi-Newtonovské metody, jenž zmíněnou složitost obcházejí aproximací Hessovy matice pozitivně definitní maticí  $\mathbf{B}_k$  (v kroku  $k$ ). Nejpoužívanější z těchto metod je BFGS. V [32] je dokázáno, že vhodná pozměňovací forma matice je tvaru (při definování jistých omezení, které zajistí jedinečnost  $\mathbf{B}$ )

$$\mathbf{B}_{k+1} = \mathbf{B}_k + a\vec{u}\vec{u}^T + b\vec{v}\vec{v}^T, \quad (3.12)$$

kde  $\vec{u}$ ,  $\vec{v}$  jsou vzájemně lineárně nezávislé vektory, respektive  $a$ ,  $b$  reálná čísla. Jinými slovy pokud je  $\mathbf{B}_0$  pozitivně definitní, tak musí být i  $\mathbf{B}_{k+1}$ , neboť hodnota matice  $\vec{u}\vec{u}^T$ , respektive  $\vec{v}\vec{v}^T$  je 1 [31].

Jednou z podmínek kladených na  $\mathbf{B}_{k+1}$  u všech kvazi-Newtonovských metod je nutná platnost

$$\mathbf{B}_{k+1}\Delta\vec{x}_k = \vec{y}_k, \quad (3.13)$$

kde  $\vec{y}_k = \nabla f(\vec{x}_{k+1}) - \nabla f(\vec{x}_k)$ . Potom je možné z (3.12) a přirozené volbě  $\vec{u} = \vec{y}_k$ ,  $\vec{v} = \mathbf{B}_k\Delta\vec{x}_k$ , získat vyjádření

$$a = \frac{1}{\vec{y}_k^T \Delta\vec{x}_k}, \quad b = -\frac{1}{\Delta\vec{x}_k^T \mathbf{B}_k \Delta\vec{x}_k} \quad (3.14)$$

Dosazením zpět do (3.12) je obdržen vztah

$$\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{\vec{y}_k \vec{y}_k^T}{\vec{y}_k^T \Delta\vec{x}_k} - \frac{\mathbf{B}_k \Delta\vec{x}_k \Delta\vec{x}_k^T \mathbf{B}_k}{\Delta\vec{x}_k^T \mathbf{B}_k \Delta\vec{x}_k} \quad (3.15)$$

Avšak z (3.13) plyne nutnost vyjádření inverzní matice  $\mathbf{B}_k$ . Například při aplikaci *Woodburovy maticové identity* (důkaz například v [31]) a přímými manipulacemi s

výrazem výše je získáno

$$\mathbf{B}_{k+1}^{-1} = \left( \mathbf{I} - \frac{\Delta \vec{x}_k \vec{y}_k^T}{\vec{y}_k^T \Delta \vec{x}_k} \right) \mathbf{B}_k^{-1} \left( \mathbf{I} - \frac{\vec{y}_k \Delta \vec{x}_k^T}{\vec{y}_k^T \Delta \vec{x}_k} \right) + \frac{\vec{y}_k \Delta \vec{x}_k^T}{\vec{y}_k^T \Delta \vec{x}_k} \quad (3.16)$$

Je možné stále dokázat, že se jedná o pozitivně definitní matici [32]. Z tohoto důvodu také předpokládejme, že funkce  $\mathcal{J}(\vec{p})$  je konvexní alespoň v dostatečně velkém okolí počátečního odhadu  $\vec{p}_0$ , neboť v opačeném případě by mohlo docházet ke konvergenci k falešným minimům jako jsou sedlové body funkce.

Algoritmus byl implementován v prostředí MATLAB. Z (3.13) je zřejmé, že každý krok vyžaduje výpočet  $\vec{y}_k$ , avšak ten je neproveditelný bez prvotního odhadu  $\vec{x}_{k+1}$ . To bylo učiněno nalezením koeficientu (hledání bylo vždy započato od hodnoty 0 až došlo ke splnění podmínky) splňující takzvanou *Wolfovu podmínku* (viz [32]).

Aby bylo zajištěno, že bude docházet ke konvergenci k tíženému minimu  $\mathcal{J}(\vec{p})$ , bylo na základě vyhodnocení určitých společných rysů některých parametrů ve článcích [21], [22], [18], [29] s parametry (3.1) a předběžného odhadu pravděpodobností popsaných v první sekci této kapitoly, určeno pro všechny kalibrace počáteční  $\vec{p}_0$  hodnoty  $\beta_P = 0.5$ ,  $\beta_I = 0.1$ ,  $\beta_H = 0.01$ ,  $\kappa = 0.3$ ,  $\gamma = 0.05$ ,  $\theta_H = 0.1$ ,  $\theta_K = 0.1$ ,  $\alpha = 0.15$ ,  $\alpha_H = 0.05$ ,  $\alpha_K = 0.05$ .  $\mu$  bylo určeno na pevnou hodnotu bez kalibrace na 0.0066, odpovídající přibližně pěti měsícům.  $m$  bylo určeno individuálně, protože může nabývat jen kladných celočíselných hodnot. Jak metrika úspěšnosti byla zvolena průměrná absolutní chyba (MAE). Běh kalibrace parametrů přes BFGS byl omezen na 100 iterací. Vždy bylo zvoleno  $\mathbf{B}_0 = \mathbf{I}$ , tedy jednotková matice. Ukázku chování MAE jako zavislosti na počtu iterací BFGS při kalibraci u hodnot pro Lombardii je na obr. 3.4 (zobrazeno pouze prvních 50 iterací). Vidíme velmi rychlou konvergenci.

## 3.4 Výsledky a diskuze

Na základě dat popsaných v sekci 3.2 a metody odhadu parametrů ze sekce 3.3 byly určeny parametry u všech zmíněných regionů a celé Itálie pro vymezené fáze a období. Jako počáteční podmínky byly zvoleny hodnoty z [30] odpovídají prvním dnům po začátku dané fáze. U kompartmentu  $P$  pro nějž neexistují data byl u první fáze proveden odhad počáteční hodnoty jako pětinasobek (určeno podle [33]) detekovaných jedinců v tomto okamžiku. V dalších fázích byl tento poměr odhadnut jako průměrný poměr hodnot mezi  $P$  a  $I$  z předchozí fáze. Podobně bylo postupováno u skupiny  $R$ , protože bylo očekáváno násobně vyšší číslo než udávané. Výsledky kalibrace parametrů jsou uvedeny v tab. 3.1. Není předložen koeficient  $c_P$ , který byl určen u všech regionů na  $\approx 0.0108$ . Velikost  $\beta_H$  vždy činila méně než zvolená hladina přesnosti ( $< 10^{-4}$ ), proto bylo stanoveno  $\approx 0$  u všech regionů. Očividně tento

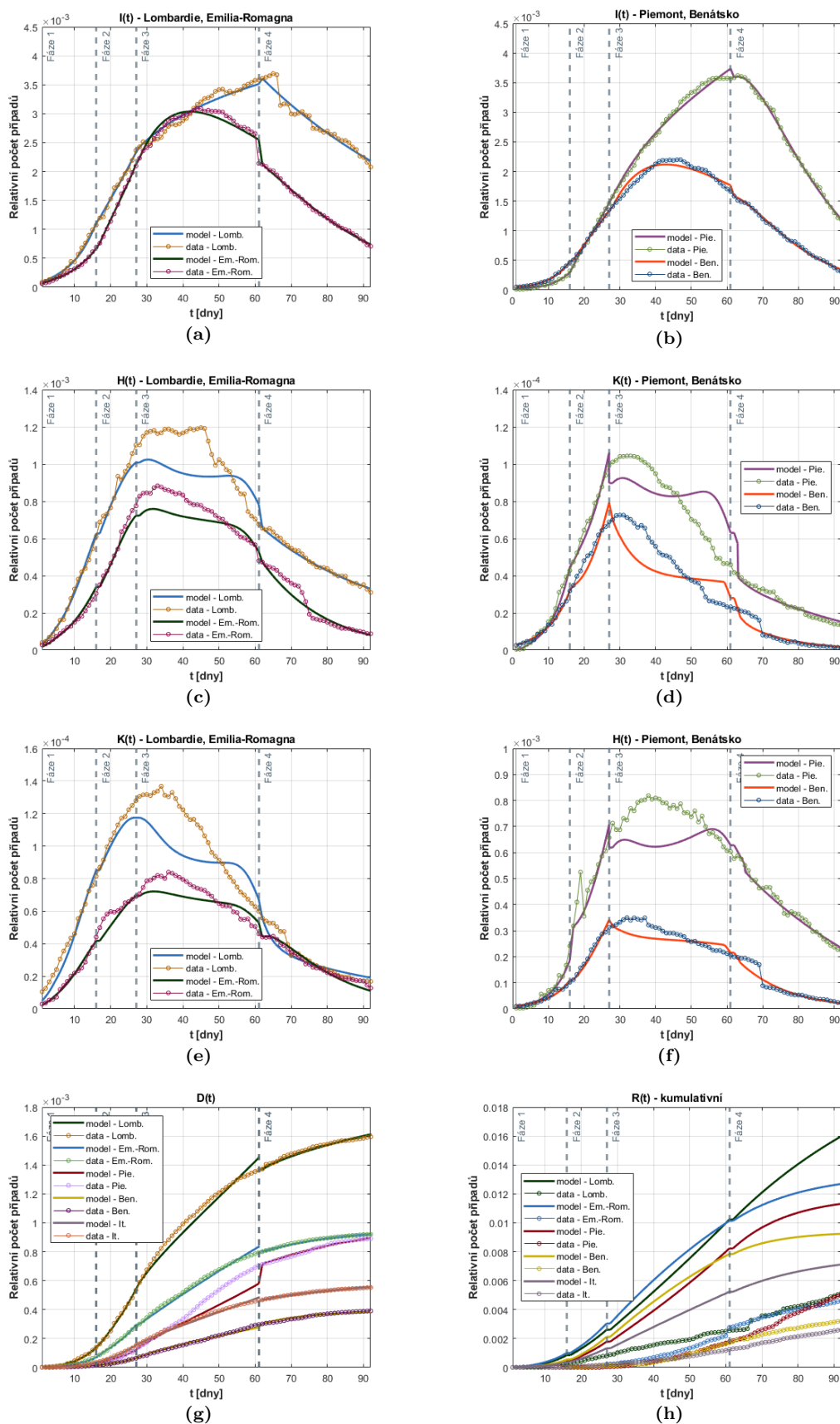


Emilia-Romagna										
Fáze	$\beta_P$	$\kappa$	$\gamma$	$\theta_H$	$\theta_K$	$\alpha$	$\alpha_H$	$\alpha_K$	$\mathcal{R}_0$	MAE
1	0.3742	0.1760	0.2520	0.0534	0.0683	0.0728	0.0386	0.0751	1.4973	4.4616
2	0.2421	0.1075	0.3529	0.0456	0.0512				1.3345	7.4862
3	0.1084	0.0509	0.2394	0.0132	0.0299				0.8685	71.9904
4	0.0813	0.0260	0.1450	0.0089	0.0225				0.8229	17.8368
Lombardie										
Fáze	$\beta_P$	$\kappa$	$\gamma$	$\theta_H$	$\theta_K$	$\alpha$	$\alpha_H$	$\alpha_K$	$\mathcal{R}_0$	MAE
1	0.4016	0.1819	0.2330	0.0959	0.0723	0.0720	0.0306	0.0717	1.5699	32.4768
2	0.2029	0.0393	0.3452	0.0695	0.0588				1.3437	17.9683
3	0.1222	0.0790	0.2725	0.0188	0.0337				1.0793	174.5813
4	0.0801	0.0443	0.2741	0.0044	0.0192				0.8822	47.2750
Piemont										
Fáze	$\beta_P$	$\kappa$	$\gamma$	$\theta_H$	$\theta_K$	$\alpha$	$\alpha_H$	$\alpha_K$	$\mathcal{R}_0$	MAE
1	0.6704	0.3956	0.2632	0.0629	0.7823	0.0771	0.0358	0.0786	1.4278	32.2200
2	0.2846	0.1487	0.1248	0.0783	0.0438				1.2784	8.3019
3	0.1426	0.0550	0.1520	0.0144	0.0301				1.1061	100.4497
4	0.0985	0.0282	0.2400	0.0158	0.0186				0.9752	19.4485
Benátsko										
Fáze	$\beta_P$	$\kappa$	$\gamma$	$\theta_H$	$\theta_K$	$\alpha$	$\alpha_H$	$\alpha_K$	$\mathcal{R}_0$	MAE
1	0.3181	0.0678	0.0771	0.0666	0.0930	0.0701	0.0342	0.0720	2.2446	2.5936
2	0.2232	0.0988	0.0930	0.0462	0.0623				1.2923	4.0740
3	0.1095	0.0562	0.1415	0.0174	0.0319				0.8415	52.3959
4	0.1002	0.0492	0.3769	0.0009	0.0319				0.8399	15.8357
Itálie										
Fáze	$\beta_P$	$\kappa$	$\gamma$	$\theta_H$	$\theta_K$	$\alpha$	$\alpha_H$	$\alpha_K$	$\mathcal{R}_0$	MAE
1	0.4023	0.1660	0.1838	0.0714	0.0791	0.0716	0.0393	0.0760	1.6768	38.6993
2	0.2309	0.1013	0.2317	0.0684	0.0587				1.3178	44.1232
3	0.1203	0.0492	0.2066	0.0146	0.0319				0.9771	381.9041
4	0.0921	0.0412	0.2069	0.0027	0.0185				0.8165	129.9535

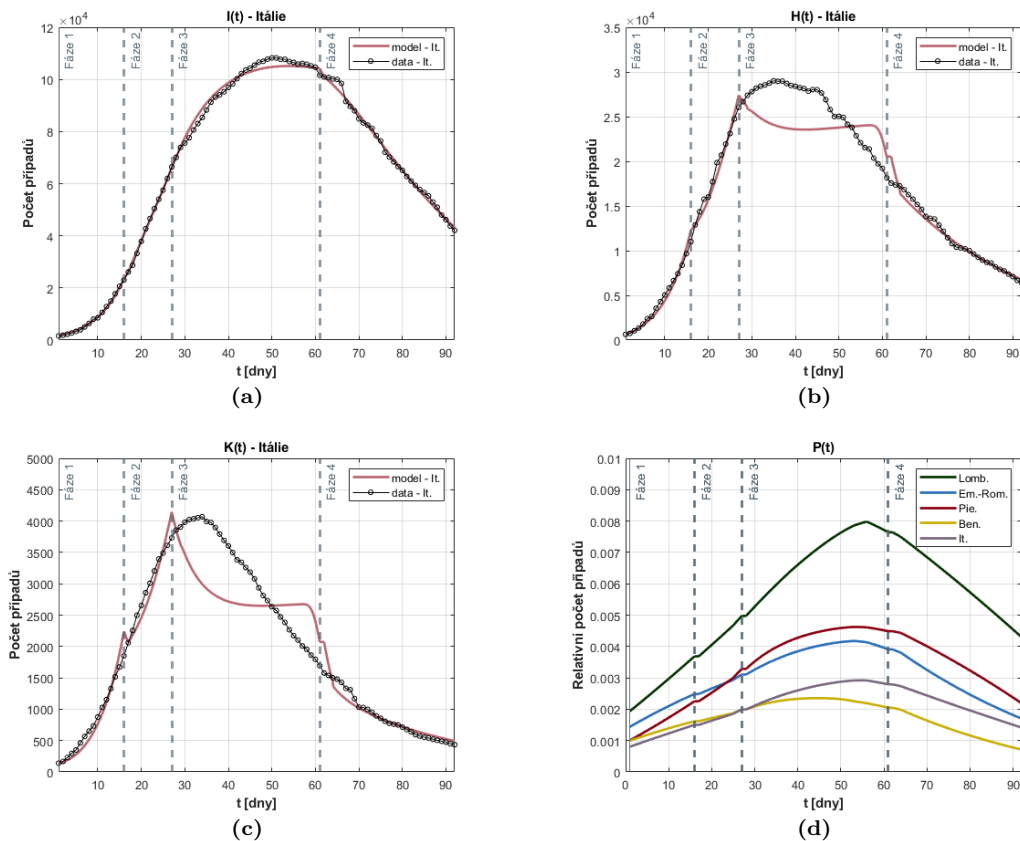
**Tab. 3.1:** Výsledky odhadu některých parametrů, MAE a  $\mathcal{R}_0$ .

fakt neodpovídá realitě, avšak pro studování tohoto fenoménu je nutné zajistit více informací například trasováním pro zdravotnický personál a využít grafové modely. Tvarový parametr  $m$  byl určen individuálně, kdy po kalibraci všech parametrů, při niž bylo dáno  $m = 3$ , bylo testována zda dochází ke snižování MAE pro různá  $m$ , ale již daných parametrech. Došlo k poklesu o jednoty až po  $m = 5$  poté již méně, tudíž tato hodnota byla určena jako závazná pro všechny regiony. Je však nutné připomenout, že koeficient  $m$  by měl být hlavně stanoven biochemickými rozbory pro danou chorobu a skupinu nemocných.

Po nahlédnutí do tabulky výše, na obr. 3.5 a obr. 3.6 lze učinit některé obecné závěry o získaných výsledcích platné pro všechny regiony. Ve všech fázích docházelo



**Obr. 3.5:** Model vs data severní Itálie. Lombardie + Emilia-Romagna: (a) detekování - mírný průběh. (c) Hospitalizovaní - nekritický stav. (e) nemocní na JIP. Piemont + Benátsko: (b) detekování - mírný průběh. (d) Hospitalizovaní - nekritický stav. (f) nemocní na JIP. (g) zemřelý (všechny regiony). (h) uzdravení (všechny regiony).



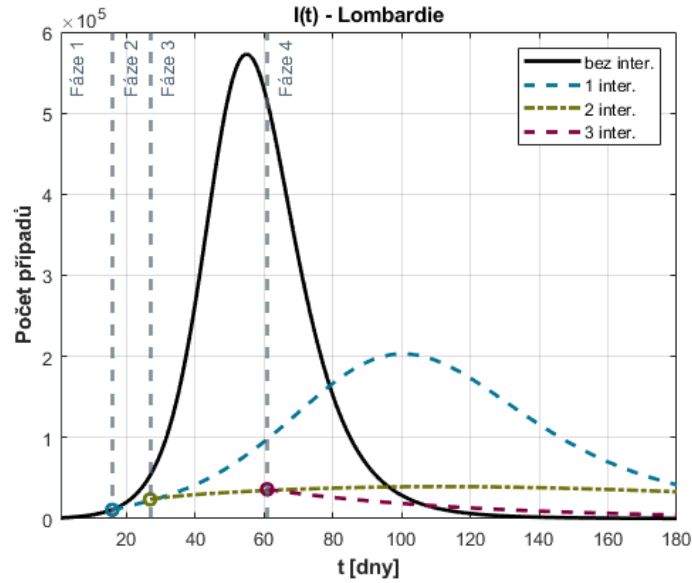
**Obr. 3.6:** Model vs data Itálie: (a) detekování - mírný průběh. (b) hospitalizování - nekritický stav. (c) nemocní na JIP. (d) nedetekování pro všechny regiony + Itálie.

k postupnému poklesu  $\beta_P$ , který odráží úspěšnost protiepidemických opatření.  $\kappa$  vykazovala setrvalý sestup (kromě fáze 1 a 2 u Benátska). To implikuje pokles pravděpodobnosti, kterou nemocný jedinec má, že bude detekován. Například u Lombardie ze 71% ve fázi 1 na pouhých 38% ve fázi 4. To je možné například reflektovat či připsat změně věkové struktury nemocných během epidemie, tj. možný nárůst asymptotických nakažených, saturaci testovacích kapacit, nárůst testování i zdravých jedinců, změna chování, kdy jedinec raději volí cestu netestování, aby nebyl komunitně ostrakizován apod. Progresivní úbytek v hodnotách je pozorovatelný i u  $\theta_H$ . Zde však lze zpochybňovat věrohodnost predikcí modelu, neboť hodnoty zejména u fází 1 či 2 naznačují, že existovala šance až 50%, že bude detekovaný jedinec hospitalizovaný. Avšak například podle [28] byla tato pravděpodobnost okolo 20% před kulminací epidemie a postupně klesající až na úroveň přibližně 1%. Do tohoto intervalu padají parametry až ve fázích 3 a 4.  $\theta_K$  naznačuje například u Benátska v první fázi až 70% pravděpodobnost přesunu na JIP. Až v pozdějších fázích u všech regionů dochází k přesunu tohoto čísla do intervalů 20% - 35%, které podle [28] jsou realistická.  $\gamma$  kolísá okolo podobných hodnot pro všechny regiony, tj. šance přežití pacienta po přesunu na JIP činí 40% - 60% (věrohodné [28]). To také naznačuje,

že věkové skupiny v  $K$  nebyly s postupující epidemií příliš diverzifikované, neboť dá se očekávat vyšší pravděpodobnost přežití u mladších osob.  $\alpha$  odpovídala přibližně průměrné době infekčnosti 14 dnů, avšak dle literatury jde spíše o 10 - 12 dnů. To je patrně způsobeno postupem jakým byl jedinec označen za zdravého a neinfekčního do systému sběru dat. Z tohoto důvodu by bylo nejspíše dobré rozdělit  $\alpha$  pro jedince z  $I$  a  $P$ , u nichž lze očekávat vyšší hodnoty parametru, v modelu (3.1). Čísla pro  $\alpha_H$  korespondují s dobou okolo 28 dní, tj. dá se očekávat, že jedinec bude déle infekční, pokud bude mít těžší průběh onemocnění.  $\alpha_K$  odpovídá také přibližně 14 dnům, avšak je nutné přičíst i dobu strávenou v  $I$ , respektive  $H$ , pro celkovou dobu infekčnosti pacienta na JIP.

V tabulce 3.1 je zjevný trend poklesu  $\mathcal{R}_0$ , tudíž v každé fázi docházelo k redukci parametrů, jenž mají dostatečně vysokou senzitivitu podle (3.7). Například ve [33] činil odhad reprodukčního čísla na počátku epidemie  $\approx 2.2$ , kdy samozřejmě záleží na použitém modelu, nicméně lze tvrdit, že model (3.1) predikuje menší závažnost počáteční fáze. Úspěšnost predikce modelu pro dané fáze je vyjádřena pomocí MAE. Velmi zdařené jsou všechny predikce pro kompartment  $I$  ve všech fázích, kde model ve všech případech vystihl dynamiku epidemie s maximálními odchylkami desítek jedinců. Obecně si model nejhůře vedl ve třetí fázi, kdy průměrná odchylka činila až více než 100 jedinců u regionů a téměř 400 u celé Itálie. To zejména z důvodu dynamiky v kompartmentech  $H$  a  $K$ . Zde model nedokáže vystihnout aktuální trend a při korekci při přechodu fází vykazuje velmi abruptní chování, jak je patrné na obr. 3.5, respektive 3.6. To může být způsobeno například špatným určením času projevu efektu kontrolních opatření dané fáze. Velmi dobrý je výsledek v kompartmentu  $D$ , kromě třetí fáze v Piemontu (při přechodu fází 3 a 4 a korekci hodnot u Lombardie a Emilia-Romagna by došlo k zdánlivému snížení počtu mrtvých, proto byl ve vyobrazení 3.5g zobrazen na tomto rozhraní nespojitě). Dynamika pro kompartmentu  $P$  na obr. 3.6d a s porovnáním s  $I$  je možné se domnívat, že počet nedetekovaných vůči počtu detekovaných činil 2 až 4. Tedy číslo je podhodnocené oproti například [33], kde autoři dovozují poměr 5 - 10. Nejlépe si model vedl při predikcích pro Benátsko, kde byla průměrná odchylka jen v jednotkách jedinců, respektive  $\approx 52$  ve třetí fázi. Nicméně je nutné přihlídnout k demografickým proměnným, kdy Benátsko je nejmenší studovaný region, a tudíž například celá Itálie vykazuje řádově mnohem vyšší odchylky. Celkově lze tvrdit, že si model 3.1 vedl lépe než jednodušší modely například v [29] či [33] a srovnatelně s komplexnějším modelem například v [28].

Pro testování efektu protiepidemických opatření byly simulovány průběhy epidemie, kdy vždy byla postupně vynechána jedna z fází kontrolních opatření. Jinými slovy bylo předpokládáno, že po dané fázi již nedošlo k žádné další intervenci. Poté byla vypočtena odhadované zasažení populace daného regionu (po poklesu pod 100 infekčních jedinců). Jak je patrné z obr. 3.7, respektive tab. 3.2 Při absenci vládních



**Obr. 3.7:** Hypotetické průběhy epidemie při absenci některé z vládních intervencí (kompartment  $I$ ).

Region	Zasažení [%]			
	bez inter.	1 inter.	2 inter.	3 inter.
<b>Emilia-Romagna</b>	61.93	51.45	4.44	3.49
<b>Lombardie</b>	65.89	50.96	16.41	4.02
<b>Piemont</b>	54.42	44.92	16.41	3.21
<b>Benátsko</b>	82.62	46.35	2.77	1.80
<b>Itálie</b>	71.54	48.85	4.76	4.31

**Tab. 3.2:** Odhad zasažení při vynechání některé vládní intervence.

opatření po 1. březnu a konstantních parametrech po celou dobu pandemie (nepravděpodobné) by došlo k extrémnímu nárůstu případů a pravděpodobně extrémě rychlé saturaci zdravotnického systému. U dalších fází je již viditelná "*flattening the curve*" strategie. Je zřejmé, že toto zploštění křivky je způsobeno poklesem v reprodukčním čísle. Test naznačuje, že až po druhé intervenci dochází k signifikantnímu poklesu konečného zasažení. Ve [28] je uveden odhad zasažení pro celou Itálii po první vlně pandemii na 4.5%, tedy jde o nepatrnou odchylku od predikce (3.1) (viz tab. 3.2).

Limitace modelu (3.1) tvoří nezahrnutí věkové struktury obyvatelstva (tj. zejména heterogenita v počtu kontaktu apod.), kdy by bylo nutné provést paralelizaci modelu dle demografického rozložení, dále diverzifikace gama rozložení pro více kompartmentů a různými tvarovými parametry či začlenění efektu mobility obyvatelstva. Při každé epidemii dojde k určité změně vzorců chování obyvatelstva, které je promítnuto zejména do koeficientů  $\beta$  a které není možné spojit s konkrétními oficiálními kontrolními opatřeními, avšak vždy je tento efekt přítomen [4].

# Závěr

Cílem této práce bylo seznámit se se specifickými matematickým modelováním v epidemiologii a na tomto základě vytvořit model, jenž by měl odrážet charakteristiky virového onemocnění COVID-19, a provést, respektive vyhodnotit úspěšnost modelu na dostupných datech ze světové pandemie COVID-19 z let 2020/2021.

V první kapitole byly představeny základní spojité deterministické kompartmentové modely jako SIS, SIR, SEIR, SEQIHR. Tyto modely byly zobecněny pro libovolné pravděpodobnostní rozdělení řídící setrvání jedinců v kompartmentu. Poté byla zkoumána rozdílnost v gama a exponenciální distribuci v jednotlivých skupinách u modelu SEQIHR. Byly dokázány signifikantní odlišnosti zejména v odhadech konečného zasažení populace a reprodukčních číslech. Bylo spekulováno obecně o větší vhodnosti gama rozdělení, neboť exponenciální rozdělení vykazuje bezpaměťový charakter. Tedy je vždy nutné zvážit použití daného modelu, protože může ovlivnit vytváření protiepidemických opatření. Dále byla numericky testováno chování některých modelů s časově proměnnými parametry. Minoritně byly představeny modely vycházející z teorie grafů a jejich uplatnění v počátku epidemie. Bylo připsána ekvivalence reprodukčního čísla a průměrného přebytkového stupně uzlu grafu. Bylo ukázáno na příkladech, že i když je tento stupeň ostře větší než 1, tak vždy existuje pravděpodobnost, že nedojde k propuknutí nákazy. Nakonec byly definovány podmínky, kdy dochází ke konvergenci předpovědí grafových a spojitých kompartmentových modelů.

Druhá kapitola byla rešerší vybraných studií ve vztahu k modelování COVID-19, především s ohledem na použitý model, analýzu dynamiky modelu, daný region, některých netriviálních poznatků plynoucích ze simulací apod.

Ve třetí kapitole byl konstruován model, který měl na základě sledovaných kategorií nemocných přístupných ve veřejné databázi pro severní Itálii první poloviny roku 2020, co možná nejvíce umožnit kalibraci vůči známým monitorovaným skupinám v databázi. To umožnilo co nejpřesněji odhadnout chování kompartmentu nedetekovaných ( $P$ ), jenž je klíčový pro dynamiku epidemie. Formulace modelu obsahovala gama ( $P, I$ ) i exponenciální distribuci ve skupinách. Parametry modelu byly určeny přes hledání minima sumy čtverců reziduí přes algoritmus BFGS. Výsledky simulací byli velmi dobré zejména u skupin  $I$  a  $D$  u všech zvolených fází kontrolních opatření, překonávající studie využívající jednodušší modely, avšak srovnatelné s modely s komplexnější strukturou. V kategorii  $H$ , respektive  $K$  docházelo zejména ve třetí fázi k divergenci modelu a reálných hodnot. Byla prokázána závěry některých studií s predikcí až trojnásobného počtu nedetekovaných vůči detekovaným jedincům. Nakonec byla provedena simulace vlivu protiepidemických opatření. Absence kontrolních opatření by vedla k extrémnímu nárůstu konečného zasažení populace.

# Literatura

- [1] Oldstone, Michael B.A. *Viruses, Plagues, and History: Past, Present and Future*. 2. vyd. New York, NY, USA: Oxford University Press, 2009. 383 s. ISBN 978-0-19-532731-1.
- [2] Higham, J. et al. *The Princeton Companion to Applied Mathematics*. Princeton, NJ, USA: Princeton University Press, 2015. 1032 s. ISBN 978-0-691-15039-0.
- [3] Kermack, W. O.; McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London*, 1927, vol. 115, no. 722, s. 700–721. ISSN 0950-1207.
- [4] Brauer, F. et al. *Mathematical Models in Epidemiology*. New York, NY, USA: Springer, 2019. 625 s. ISBN 978-1-4939-9826-5.
- [5] Brauer, F. et al. *Mathematical Models in Population Biology and Epidemiology*. New York, NY, USA: Springer, 2012. 522 s. ISBN 978-1-4614-1685-25.
- [6] Brauer, F. et al. *Mathematical Epidemiology*. New York, NY, USA: Springer, 2008. 412 s. ISBN 978-3-540-78910-9.
- [7] Coddington, E.A.; Levison, N. *Theory of Ordinary Differential Equations*. Malabar, FL, USA: Krieger Publishing, 1984. 429 s. ISBN 978-0898747553.
- [8] Newman, M.E.J. The spread of epidemic disease on networks. *Physical Review E*, 2002, vol. 66, no. 016128, s. 1–11. ISSN 2470-0053.
- [9] Dietz, K. Overall patterns in the transmission cycle of infectious disease agents. *Report of the Dahlem Workshop on Population Biology of Infectious Disease Agents Berlin 1982*, 1982, vol. 25, s. 87–102. ISBN 978-3-642-68637-5.
- [10] Feng, Z. et al. Epidemiological Models with Non-Exponentially Distributed Disease Stages and Applications to Disease Control. *Bulletin of Mathematical Biology*, 2007, vol. 69, s. 1511—1536. ISSN 1522-9602.
- [11] Casella, G.; Berger, R.L. *Statistical Inference*. 2. vyd. Pacific Grove, CA, USA: Duxbury, 2002. 688 s. ISBN 0-534-24312-6.
- [12] Blitzstein, J.K.; Hwang, J. *Introduction to Probability*. Boca Raton, FL, USA: CRC Press, 2008. 631 s. ISBN 978-1-1383-6991-7.
- [13] Axler, S.J. *Linear Algebra Done Right*. New York, NY, USA: Springer, 1995. 250 s. ISBN 978-3-319-11080-6.

- [14] Strogatz, S. *Nonlinear dynamics and Chaos: With applications to physics, biology, chemistry, and engineering*. Boston, MA, USA: Addison-Wesley Publishing, 2014. 528 s. ISBN 978-0-8133-4910-7.
- [15] Small, M. & Tse, C. Modeling the SARS Outbreak in Hong Kong with Small World or Scale Free Networks. *International Journal of Bifurcation and Chaos*, 2005, vol. 8, s. 1745–1755. ISSN 0218-1274.
- [16] Chowell, G. et al. Characterizing the Epidemiology of the 2009 Influenza A/H1N1 Pandemic in Mexico. *PLOS Medicine*, 2011, vol. 15, s. 1–12. ISSN 1549-1676.
- [17] De Brouwer, E. et al. Modeling the COVID-19 outbreaks and the effectiveness of the containment measures adopted across countries. *medRxiv* [online]. 2020, [cit. 04. 04. 2021]. Dostupné z URL: <<https://doi.org/10.1101/2020.04.02.20046375>>
- [18] He, S. et al. SEIR modeling of the COVID-19 and its dynamics. *Nonlinear Dynamics*, 2020, vol. 101, s. 1667—1680. ISSN 1756-0500.
- [19] Skokos, Ch. et al. *Chaos Detection and Predictability*. New York, NY, USA: Springer, 2016. 269 s. ISBN 978-3-662-48408-1.
- [20] Oz, Y. et al. Heterogeneity and Superspreading Effect on Herd Immunity. *arXiv* [online]. 2020, [cit. 23. 03. 2021]. Dostupné z URL: <<https://arxiv.org/abs/2009.01923>>
- [21] Biswas, S. et al. COVID-19 pandemic in India: a mathematical model study. *Nonlinear Dynamics*, 2020, vol. 102, s. 537—553. ISSN 0924-090X.
- [22] Garba, S. et al. Modeling the transmission dynamics of the COVID-19 Pandemic in South Africa. *Mathematical Biosciences*, 2020, vol. 328, s. 1–35. ISSN 0025-5564.
- [23] *COVID-19 South African Coronavirus News and Information*. [online] 2021, [cit. 17. 04. 2021]. Dostupné z URL: <<https://sacoronavirus.co.za>>
- [24] Yang, H. et al. Mathematical model describing CoViD-19 in São Paulo, Brazil – evaluating isolation as control mechanism and forecasting epidemiological scenarios of release. *Epidemiology and Infection*, 2020, vol. 148, s. 1–22. ISSN 0253-9624.
- [25] Chitnis, N. et al. Determining important parameters in the spread of malaria through the sensitivity analysis of a mathematical model. *Bulletin of Mathematical Biology*, 2008, vol. 70, s. 1272–1296. ISSN 0092-8240.



- [26] Berman, A. & Plemmons, R. *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1994. 361 s. ISBN 0-89871-321-8.
- [27] *Coronavirus in Italy: Timeline of the COVID-19 Pandemic in Italy 2020*. [online] 2020, [cit. 25. 04. 2021]. Dostupné z URL: <<https://www.italofile.com/coronavirus-in-italy>>
- [28] Giordano, G. et al. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, 2020, vol. 26, s. 1–32. ISSN 1078-8956.
- [29] Piccolomini, E. & Zama, F. Monitoring Italian COVID-19 spread by a forced SEIRD model. *PLOS ONE*, 2020, vol. 15, s. 1–17. ISSN 1932-6203.
- [30] *Presidenza del Consiglio dei Ministri, Dipartimento della Protezione Civile, Italia: COVID-19 Italia - Monitoraggio situazione*. [online] 2021, [cit. 02. 05. 2021]. Dostupné z URL: <<https://github.com/pcm-dpc/COVID-19>>
- [31] Artin, M. *Algebra*. Harlow, Essex, UK: Pearson, 1991. 486 s. ISBN 978-1-292-02766-1.
- [32] Fletcher, R. *Practical methods of optimization*. Chichester, NY, USA: Wiley, 1987. 436 s. ISBN 0-471-49463-1.
- [33] Quaranta, G. et al. Understanding COVID-19 nonlinear multi-scale dynamic spreading in Italy. *Nonlinear Dynamics*, 2020, vol. 101, s. 1583–1619. ISSN 1573-269X.